colección TexTos 1

MATEMÁTICAS BÁSICAS PARA ECONOMISTAS

ÁLGEBRA LINEAL

con notas históricas y contextos económicos

SERGIO MONSALVE

Editor

UNIVERSIDAD NACIONAL DE COLOMBIA

VICERRECTORÍA ACADÉMICA

EDITORIAL

FACULTAD DE CIENCIAS ECONÓMICAS

MATEMÁTICAS BÁSICAS PARA ECONOMISTAS

VOLUMEN 1

ÁLGEBRA LINEAL

MATEMÁTICAS BÁSICAS PARA ECONOMISTAS 1

ÁLGEBRA LINEAL

Con notas históricas y contextos económicos

SERGIO MONSALVE

EDITOR

FACULTAD DE CIENCIAS ECONÓMICAS

UNIVERSIDAD NACIONAL DE COLOMBIA

Catalogación en la publicación Universidad Nacional de Colombia

Matemáticas básicas para economistas: con notas históricas y contextos económicos / ed. Sergio Monsalve. - Bogotá : Universidad Nacional de Colombia. Facultad de Ciencias Económicas, 2009
4 v.

Incluye referencias bibliográficas

Contenido : v. 0. Fundamentos. – v. 1. Algebra lineal. – v. 2. Cálculo. – v. 3. Optimización y dinámica

ISBN 978-958-719-304-6 (v. 0). - ISBN 978-958-719-305-3 (v. 1). -
ISBN 978-958-719-306-0 (v. 2). - ISBN 978-958-719-307-7 (v. 3)

1. Matemáticas 2. Modelos económicos 3. Matemáticas para economistas
4. Álgebra lineal 5. Cálculo 6. Optimización matemática 7. Programación dinámica

I. Monsalve Gómez, Sergio, 1962-, ed.

CDD-21 510.2433 / 2009

Matemáticas Básicas para
Economistas 1: Álgebra Lineal

Primera edición, 2009
ISBN: 978-958-719-305-3

Diseño de carátula
Ángela Pilone Herrera

Corrección de estilo
César Cortés
Ana Patricia Tolosa

Diseño de páginas interiores y
armada electrónica
Nathalie Jiménez Millán

Colaboradores del autor:
Olga Manrique
Escuela de Economía
Universidad Nacional de Colombia,
Bogotá

Francisco Lozano
Escuela de Economía
Universidad Nacional de Colombia,
Bogotá

Índice general

1. Lección 1
Sistemas de ecuaciones lineales: solución por eliminación gaussiana 1

1. Sistemas de ecuaciones lineales 1
2. Método de eliminación gaussiana 3
   a. Algoritmo de eliminación gaussiana 4
   b. Una visión geométrica 10
3. Contexto económico 18
   a. Sobre el álgebra lineal en la teoría económica 18

2. Lección 2
Matrices y determinantes 29

1. La noción de matriz 30
2. Tipos de matrices 32
3. Álgebra de matrices 34
   a. Suma de matrices 34
   b. Multiplicación de un escalar por una matriz 36
   c. Multiplicación de matrices 38
4. Otros tipos de matrices 47
   a. Matrices particionadas 58
5. Determinante de una matriz cuadrada 62
   a. Determinantes 2×2 63
   b. Determinantes 3×3 64
   c. Determinantes $n \times n$ 68
6. Propiedades de los determinantes 73
7. Contexto económico 84
   a. Primer modelo lineal formal en la teoría económica: sobre las tasas de intercambio (Cournot (1838)) 84

3. Lección 3
Sistemas de ecuaciones lineales: solución por matriz inversa **93**
1. La matriz inversa . . . 94
2. Cálculo de la matriz inversa mediante el método gaussiano . . . 100
3. Cálculo de la matriz inversa mediante determinantes (regla de Cramer) . . . 108
   a. Determinantes de matrices particionadas . . . 116
   b. Inversas de matrices particionadas . . . 117
4. Contexto económico . . . 121
   a. Una “visión lineal” en la teoría del valor: la teoría de la imputación de von Wieser (1889) . . . 121

4. Lección 4
Vectores **131**
1. El concepto de vector . . . 132
2. Norma de un vector en $\mathbb{R}^n$ . . . 139
3. Ángulo entre vectores . . . 143
   a. Proyección de un vector sobre otro . . . 150
   b. Producto cruz de vectores . . . 151
4. Rectas y planos . . . 157
   a. Rectas en $\mathbb{R}^n$ . . . 157
   b. Planos en $\mathbb{R}^n$ . . . 161
5. Contexto económico . . . 166
   a. El modelo de equilibrio general Walras-Cassel (1918) . . . 166

5. Lección 5
Bases y dimensión **179**
1. Definición de espacio vectorial . . . 180
   a. Combinaciones lineales . . . 186
   b. Subespacios vectoriales . . . 189
2. Las nociones de base y dimensión . . . 196
   a. Dependencia e independencia lineal . . . 196
3. Bases ortonormales para $\mathbb{R}^{\mathbf{n}}$ . . . 209
4. Bases para el espacio-solución de un sistema de ecuaciones lineales homogéneo . . . 215
5. Contexto económico . . . 219
   a. El análisis insumo-producto de Leontief (1936) . . . 219

6. Lección 6
Transformaciones lineales **235**
1. Transformaciones lineales . . . 240

a. Transformaciones ortogonales 247
2. Núcleo e imagen: dos subespacios asociados a una transformación lineal 251
3. Transformaciones lineales y matrices 258
a. El rango de una matriz 264
4. Estructura de los conjuntos de transformaciones lineales 272
5. Isomorfismos 277
6. Contexto económico 281
a. El modelo de equilibrio general de Von Neumann (1932) 281

7. Lección 7
Diagonalización en $\mathbb{R}^n$ 293
1. Valores propios y vectores propios de una transformación lineal 293
2. Diagonalización 299
3. Diagonalización de matrices simétricas: el teorema espectral . . 307
4. Formas cuadráticas 310
5. Breve nota sobre la diagonalización en bloques de Jordan . . . 319
6. Contexto económico 323
a. El modelo teórico de Sraffa (1960) 323

8. Lección 8
Conjuntos convexos 341
1. Noción de conjunto convexo 341
2. Introducción a la programación lineal 350
3. Contexto económico 356
a. Sobre la noción de convexidad en economía 356
b. Tres modelos lineales básicos de la teoría económica . . 357

Bibliografía 387

Respuestas 409

Índice alfabético 431

La ciencia se ha construido para satisfacer
ciertas necesidades de nuestra mente;
ella nos describe.
Y aunque tiene cierta relación con el mundo real,
esa relación es muy, muy compleja.

Robert J. Aumann
Premio Nobel de Economía 2005

Sergio Monsalve le dedica este esfuerzo a
su profesor de matemáticas Jairo Charris

A la memoria de Juan Alonso, Jorge Diego, Nancy y Adriana

Presentación general

Este libro es el resultado de varios años de trabajo de los autores como profesores de matemáticas y/o economía para las Facultades de Ciencias y Ciencias Económicas de las universidades Nacional (sedes Medellín y Bogotá), Externado de Colombia y Pontificia Javeriana, y su objetivo central es exponer algunos de los elementos fundamentales del lenguaje matemático que deberían ser comunes a todos los estudiantes de economía de nuestras épocas. Pensando en esto, hemos optado por escribir el texto en cuatro volúmenes: en el volumen 0 (Fundamentos) presentamos los requisitos matemáticos que el estudiante debe llenar para acceder más cómodamente al corpus total; el volumen 1 consiste en las nociones básicas del álgebra lineal; el volumen 2 en las nociones básicas del cálculo diferencial e integral, y el volumen 3 en las nociones básicas de la teoría de la optimización y de la dinámica.

En cada uno de los cuatro volúmenes hemos dividido los temas tratados a través de *lecciones* con un *tratamiento matemático riguroso y sin referencia a aplicación económica alguna*. Todas estas lecciones presentan, además, notas históricas que esperamos ayuden a trazar el devenir de los conceptos matemáticos que se desarrollan al punto. Por tanto, aquellos que consideran que un curso de matemáticas básicas para economistas debería ser solo eso y no un curso con aplicaciones, estarán aquí servidos. Sin embargo, para aquellos que difieren de esta postura metodológica y pedagógica *hemos también separado la sección final de casi todas las lecciones para el* "contexto económico". Pero esta no es una sección ordinaria de aplicaciones a la economía: es, por el contrario, una aproximación coherente a problemas centrales en la teoría económica, y una orientación para el estudiante atento y disciplinado. Por ejemplo, en el volumen 1 aparecen discusiones sobre los modelos lineales fundamentales de la teoría económica: el modelo *walrasiano* de Cassel, el modelo *insumo-producto* de Leontief, el modelo de *equilibrio general* de von Neumann, el modelo *sraffiano*, la teoría de juegos de von Neumann y Morgenstern, el modelo "keynesiano" lineal IS-LM, y el *análisis de actividades* de Koopmans. En el volumen 2 se encuentran, entre otras discusiones, notas históricas y de contexto del problema de la racionalidad, de la revolución marginalista y de la comunión entre racionalidad y marginalismo; en el volumen 3 aparecen tres de las visiones modernas más importantes sobre el comportamiento económico: el modelo keynesiano IS-LM no-lineal de Hicks, el modelo walrasiano de Arrow y Debreu, y los modelos de interacciones económicas y sociales. El objetivo en cada uno de estos análisis es el problema económico por sí mismo y las consecuencias que el desarrollo lógico de las hipótesis y las herramientas matemáticas entregan para discusión tanto a nivel teórico-conceptual como de política económica. En ningún caso se centra en las herramientas matemáticas que están siendo utilizadas.

En definitiva, este trabajo es una invitación a comenzar a entender el potencial y, sobre todo, los límites de la herramienta matemática *tradicional* en la teoría económica; es una invitación a entender que las matemáticas *tradicionales* están mejor diseñadas y adaptadas a las ciencias exactas como la física, pero quizás no para el estudio de los fenómenos sociales y económicos, y esto intentamos resaltarlo en el texto cuando presentamos numerosos ejemplos tomados de la física, de la química, o de la biología. Pero aunque estamos convencidos de que las matemáticas son más claras que cualquier otro lenguaje y de que en numerosas ocasiones muestran lo que no podría lograrse por introspección, probablemente el verdadero aporte de ellas a las ciencias sociales y económicas únicamente podrá ser evaluado por las generaciones futuras; no antes y, por supuesto, no ahora. Solo que en ese camino no deberíamos seguir ni la moda del día, ni la aprobación o desaprobación de nuestros colegas. En su lugar, nos debería preocupar alcanzar más y más claras comprensiones de lo que sucede en los fenómenos económicos que enfrentamos día a día, y si estas u otras matemáticas son un mecanismo apropiado para lograrlo, habríamos avanzado un paso más en este propósito.

Una palabra final. Algunos tienen la creencia de que no hay manual ni texto, por bueno que sea, que pueda relevarnos de la lectura de los artículos originales y de los textos clásicos; y que nadie debería permitirse que "le cuenten" lo que dicen los escritos originales. Pero creemos que esta es una opinión, por lo menos, falaz. Claro está que es ideal poder leer los textos originales y los clásicos. Sin embargo, el estudiante que apenas se insinúa en cualquier área del conocimiento, requiere de *esquemas* y de *puntos de referencia* para poder avanzar con mayor seguridad y consistencia; posteriormente, una vez haya adquirido cierta *madurez* y *entendimiento*, es *absolutamente necesario* que recurra, ahora sí, a los textos clásicos y a los originales. Comenzando por esta estrategia, un estudiante correrá, creemos, un menor riesgo de confundirse o, lo que sería fatal, de extraviarse definitivamente.

Por último, ha sido un honor para quien esto escribe, haber podido realizar en compañía de su antiguo profesor de matemáticas de la Universidad Nacional de Colombia, sede Medellín, Fernando Puerta, los volúmenes 0 y 2 de este texto.

Agradecemos a las Facultades de Ciencias y Ciencias Económicas de la Universidad Nacional de Colombia, en particular a los profesores Carlos Andrés Álvarez (Coordinador de Publicaciones de la Facultad de Ciencias Económicas) por su inmensa disposición en el proceso de producción de este libro. También a la Facultad de Economía de la Universidad Externado de Colombia, y al Departamento de Matemáticas de esta universidad. De igual manera a aquellos de los que recibimos sugerencias y comentarios: Diego Arévalo, Julián Arévalo, Oscar Benavides, Catalina Blanco, Lina Cañas, Angélica Chappe, Lola Coba, Luis Jorge Ferro, Jorge Gallego, Norma Gómez, Carlos Augusto Jiménez, Gustavo Junca, Crescencio Huertas, Norman Maldonado, Juliana Moncada, Eduardo Mantilla, Ángela Ospina, Diego Pardo, Sergio Parra, Carolina Peláez, Lida Quintero, Aida Sofía Rivera, María Cristina Rodríguez, Diego Rojas, Marcela Rubio, Renata Samacá, Alejandra Sánchez, Humberto Sarria, Biviana Suárez, Jennifer Taborda, María del Pilar Tejada, Ana Tamayo, Héctor Useche y Miguel Zárate. Un agradecimiento del editor al Banco de la República por su apoyo en la realización de estudios de economía a nivel de doctorado (University of Wisconsin-Madison y The Hebrew University of Jerusalem). A la editorial de la Universidad Nacional de Colombia, en especial a su director Luis Ignacio Aguilar y a su jefe de editorial Dora Inés Perilla, nuestro reconocimiento por el inmenso trabajo realizado. También a Maribel Romero, Santiago Sierra, Danny Sierra, Dora Millán y Nathalie Jiménez, por su paciente digitación de nuestros difíciles manuscritos. Pero, por encima de todo, a nuestras familias que son el gran aliento y nuestra razón de ser.

Sergio Monsalve
Bogotá D.C., febrero de 2008

Nota del editor para el volumen 1

Una vez presentado el volumen 0 (*Fundamentos*) de *Matemáticas básicas para economistas*, creímos que el primer paso en la formación matemática de todo economista moderno era afrontar el estudio de aquellas herramientas que permiten abordar "problemas lineales"; es decir, de lo que hoy llamamos *álgebra lineal*. Al plantearlo así, decidimos tomar, como hilo articulador, *la solución de un sistema de ecuaciones lineales*, pues este problema, aparentemente simple, es el verdadero origen de una gran cantidad de conceptos e ideas del álgebra lineal: matriz, determinante, base, dimensión, etc.

Y aunque el tratamiento formal de este texto lo podría asemejar a otros de este mismo nivel y objetivo, se diferencia de ellos en varios aspectos: en primer lugar, en la orientación que hemos dado a la conformación de las lecciones, alrededor de los sistemas de ecuaciones lineales, y el hacerlo siempre acompañado de su respectiva *conexión geométrica*. En segundo lugar, la presentación (en los "contextos económicos" de final de cada lección), de los más importantes modelos económicos lineales, que son aún hoy estudiados en nuestras carreras de economía, así sólo sea, en algunos casos, para propósitos de fundamentación teórica (el modelo Walras-Cassel, el modelo de Leontief, el modelo de von Neumann, el modelo de Sraffa, el modelo IS-LM lineal, el modelo de juegos de von Neumann-Morgenstern, y el modelo de mercado competitivo de Koopmans); y, finalmente, las notas históricas que nos permiten, de cierta forma, trazar el devenir de los conceptos matemáticos y económicos que hemos desarrollado.

Una observación pedagógica. Este texto ha sido elaborado, no pensando exclusivamente en los estudiantes de pregrado de economía; también es apropiado para estudiantes de maestría y doctorado. Sólo que en la presentación del material, al profesor o instructor le correspondería hacer el *énfasis adecuado* que mejor se adapte al curso o seminario que tenga a cargo. Por ejemplo, es nuestra recomendación que en los cursos de pregrado la mayor parte de (aunque no todas) las demostraciones sean evitadas, y que el énfasis se haga sobre la correcta aplicación de los conceptos y teoremas. Así, creemos, el estudiante podrá hacer un mejor tránsito de la intuición al rigor.

Debemos anotar que es primordial hacer una cantidad muy apreciable de los ejercicios propuestos, pues son una parte vital del desarrollo armónico del curso. A algunos les hemos proveído respuesta, y a otros de un asterisco (*) (o dos (**)), significando esto que, quizás, tienen un nivel de dificultad un poco mayor que el resto de ejercicios. Sin embargo, es nuestra opinión que el entregar una cantidad abundante de respuestas es inconveniente, pues casi cualquier estudiante universitario tiene acceso a *softwares* tales como MATLAB, con los cuales puede corroborar muchas respuestas, en una rutina simple y, sobre todo, enriquecedora.

Finalmente, y aunque ya fue señalado en la presentación general del texto, el lector en ocasiones se podría sorprender, en este y en los siguientes volúmenes, de que en las discusiones temáticas aparezcan, con énfasis, ilustraciones y ejemplos de física, química, biología, y no de economía; sobre lo cual debemos decir que *esto está en la misma concepción del texto*, pues, al fin y al cabo, las herramientas matemáticas que hoy utilizan las ciencias económicas surgieron más de las necesidades de aquellas ciencias, y que las matemáticas en economía "llegaron tarde" en esta construcción. Por tanto, es natural que las matemáticas que presentamos en estos volúmenes estén mejor adaptadas a ellas, y por ello aparecen esos interesantes ejemplos de aplicación. No obstante, repetimos también, los "contextos económicos" al final de cada lección están dedicados a un *problema o modelo económico concreto*, acompañado de ejemplos resueltos y también propuestos, esperando con esto justificar por qué este es un texto diseñado para estudiantes de economía y, en general, para estudiantes de las ciencias económicas.

Varias advertencias de notación para los cuatro volúmenes. Los números con expresión decimal se escriben utilizando el punto (.) para separar la cantidad entera de la decimal. No se recurre a la notación, también común, de la coma (,). Utilizamos la notación $\blacksquare$ para indicar que una demostración ha finalizado, y la notación $\blacktriangle$ para indicar que un ejercicio (o ejemplo) ha terminado.

Quisiéramos creer, entonces, que la presencia de las diferencias que hemos destacado, hará de este volumen un buen aporte a las nuevas generaciones de economistas y, ojalá, también de econometristas que formamos.

Prólogo

Por: Eduardo Mantilla P.

En esta obra se recogen las experiencias didácticas de los autores en la enseñanza de la matemática, especialmente en las carreras de ciencias económicas, tomando como eje central el trabajo de varios años del profesor Sergio Monsalve.

Los textos hechos a partir de los apuntes de clase tienen el encanto de traslucir la manera de trabajar del maestro. Su aproximación a los temas. Su particular manera de decir las cosas para hacerlas comprensibles a los estudiantes. Su forma de acercarse al conocimiento. A qué le da prelación. Un texto hecho así es como una radiografía del alma pedagógica del maestro. Por eso es tan importante que no se pierdan las experiencias de quienes trabajan bien, para que otros las aprovechen e, inspirados en ellas adelanten su labor docente y cimenten su formación como educadores.

Esta obra refleja una manera de hacer las cosas de manera atractiva y rigurosa y, en cuanto a su contenido, completa para las carreras de ciencias económicas. Sus autores logran darle unidad y sabor en un trabajo dispendioso para ellos y útil para quienes tienen a su cargo asignaturas de matemáticas que aquí pueden seleccionar los temas que les sean necesarios, con la seguridad de que están bien tratados y son accesibles para los estudiantes.

Al ver la totalidad de la obra resalta el enorme trabajo que significó para el profesor Monsalve y sus compañeros recoger, ordenar y reelaborar sus experiencias y presentarlas como lo hacen. Para quien esto escribe, es especialmente atractivo el manejo de los temas geométricos que tan buenos resultados dan desde el punto de vista formativo y para la comprensión general de la materia. La presentación de modelos económicos y las notas históricas son herramientas formidables para mostrar y dar un contexto al devenir de los conceptos matemáticos y su utilización por parte de la economía.

Los autores merecen felicitaciones y el reconocimiento de la comunidad universitaria por haberse comprometido en tamaña tarea, y por la forma cuidadosa en que lo hicieron. Por lo bien que les quedó, y por lo útil que será para las futuras promociones de estudiantes. Ojalá esta obra sea probada por otros maestros que, en la práctica, son los que, con su frecuente utilización, califican la excelencia de este tipo de trabajo.

Lección 1

Sistemas de ecuaciones lineales: solución por eliminación gaussiana

Introducción

Desde la antigua Babilonia en el siglo III a. C. han aparecido, en la historia de las matemáticas, numerosos problemas resueltos a través de la solución de una o varias ecuaciones lineales; es decir, mediante operaciones simples como resolver la ecuación $ax = b$ o el sistema $ax + by = c$, $dx + ey = f$, donde a, b, c, d, e, f son números conocidos y x, y, son las incógnitas. "Pensar linealmente", es decir, con las herramientas de lo que se ha dado en llamar "álgebra lineal" ha tenido un tránsito desde la solución de estas elementales ecuaciones lineales hasta la estructuración matemática en el siglo XX, y es así como esta área llegaría a tener el lugar central que hoy ocupa en el quéhacer matemático. En los próximos ocho capítulos mostraremos cómo y por qué esto ha sido así.

1. Sistemas de ecuaciones lineales

Muchos problemas prácticos pueden plantearse como un conjunto de ecuaciones lineales. Por ejemplo, es posible que necesitemos encontrar ciertos números x, y, z tales que, simultáneamente,

$$\begin{aligned} x + 3y - z &= 1 \\ 3x - y + z &= 0 \\ x + y + z &= 2 \end{aligned} \tag{1}$$

A un conjunto de ecuaciones lineales como este se le llama un *sistema de ecuaciones lineales*, y la colección de valores que pueden tomar las variables x, y y z que satisfacen el sistema se denomina una *solución del sistema*. Así, puede verificarse, reemplazando en el sistema de ecuaciones (1), que una solución es $x = -\frac{1}{6}$, $y = \frac{5}{6}$ y $z = \frac{4}{3}$.

En forma general, tenemos la siguiente definición:

Definición 1. (Sistema de ecuaciones lineales)
Un *sistema de m ecuaciones lineales con n incógnitas* es un sistema de la forma

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad \vdots \qquad \vdots \quad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

donde los coeficientes $a_{11}, ..., a_{mn}; b_1, b_2, ..., b_m$, son conocidos.

Una solución a este sistema de ecuaciones lineales es una colección de números x_1^*, x_2^*,..., x_n^* que satisfacen, simultáneamente, las m ecuaciones.

De otro lado, este sistema de ecuaciones lineales se llamará *homogéneo* si todas las constantes b_1, b_2, $\ldots$, b_n son iguales a cero; en otro caso, el sistema se llamará *no-homogéneo*.

Nota 1.
Como decíamos antes, alrededor de los años 300 a. C., los babilonios ya estudiaban problemas que se expresarían hoy mediante sistemas de ecuaciones lineales. Posteriormente aparece la primera de las contribuciones de los griegos a la solución de un sistema de ecuaciones lineales que se debe a Diofanto (275 d. C.) quien en su *Arithmetica* utilizaba un difícil y engorroso método de solución. El problema fue extendido notablemente por los chinos, quienes utilizaban rollos de bambú para representar los coeficientes dentro de los cálculos.

Ya posteriormente, en el siglo IX, se encuentran en India un número considerable de tales problemas junto con reglas que muestran métodos para resolver ecuaciones lineales simultáneas de distintos tipos. Sin embargo, la contribución fundamental a la solución de las ecuaciones lineales la hicieron los árabes al reconocer las condiciones para la transposición de términos y la reducción a una sola incógnita; este método fue desarrollado estructuradamente por Al-Khwarizmi [780 -835] en el año 825.

Los algebristas del siglo XVI le dieron, relativamente, poca atención a las ecuaciones lineales simultáneas. De hecho, el uso de incógnitas (x, y, z ó x_1, x_2, x_3) para las cantidades desconocidas no fue sugerido sino hasta el siglo XVII. En éste y en el siglo XVIII se comenzaron a reconocer los métodos árabes para la solución de los sistemas lineales; esto debido, en gran medida, a la influencia de François Viete, René Descartes, y, particularmente, de Leonhard Euler y Karl F. Gauss. Esta discusión la ampliaremos en lecciones posteriores del texto.

Ejercicios 1

1) Según la definición 1, identifique los coeficientes a_{ij} y b_i del siguiente sistema de ecuaciones lineales:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 + x_4 &= 10 \\ 2x_1 + x_2 + 2x_3 + x_4 &= 5 \\ x_1 + 3x_2 + x_3 + 2x_4 &= 15 \\ x_1 + x_2 + x_3 + x_4 &= 12 \end{aligned}$$

2) De manera similar, identifique los coeficientes a_{ij} y b_i del sistema de ecuaciones:

$$\begin{aligned} x + 2y + z + 3w &= 10 \\ 2x + y + 3z + w &= 18 \\ 3x + y + z + 2w &= 15 \\ x + 3y + 2z + w &= 12 \end{aligned}$$

2. Método de eliminación gaussiana

Para ilustrar cómo es posible resolver, en general, los sistemas de ecuaciones lineales, regresemos al sistema (1) de tres ecuaciones con tres incógnitas

$$\begin{aligned} x + 3y - z &= 1 \\ 3x - y + z &= 0 \\ x + y + z &= 2 \end{aligned} \tag{1}$$

Allí puede llevarse a cabo un procedimiento *algorítmico* fundamental: uno puede, inicialmente, multiplicar la primera ecuación por -3 y sumar el resultado a la segunda ecuación para producir otra ecuación lineal que reemplace a la segunda del sistema (1) y así obtener el nuevo sistema lineal

$$\begin{aligned} x + 3y - z &= 1 \\ -10y + 4z &= -3 \\ x + y + z &= 2 \end{aligned} \tag{2}$$

Continuando de manera similar con el proceso, podemos ahora restar la primera ecuación a la tercera, y que la ecuación lineal resultante reemplace a la tercera ecuación del sistema (2), para obtener el sistema lineal (3):

$$\begin{aligned} x + 3y - z &= 1 \\ -10y + 4z &= -3 \\ -2y + 2z &= 1 \end{aligned} \tag{3}$$

Ahora: multiplicando por $-\frac{1}{10}$ la segunda ecuación del sistema (3) se tiene el siguiente sistema:

$$\begin{aligned} x+3y-z &= 1 \\ y-\frac{2}{5}z &= \frac{3}{10} \\ -2y+2z &= 1 \end{aligned} \tag{4}$$

En el sistema (4), multiplicamos por 2 la segunda ecuación y la sumamos a la tercera para obtener otra ecuación lineal que reemplaza la tercera ecuación del sistema, para ahora obtener el sistema lineal (5):

$$\begin{aligned} x+3y-z &= 1 \\ y-\frac{2}{5}z &= \frac{3}{10} \\ \frac{6}{5}z &= \frac{8}{5} \end{aligned} \tag{5}$$

Y ya, en este último sistema, es fácil leer la solución del mismo: $z=\frac{4}{3}$. Reemplazando "hacia atrás" este valor en la segunda ecuación del sistema se obtiene que $y=\frac{5}{6}$; y de la primera ecuación se tiene que $x=-\frac{1}{6}$. Notablemente, podemos observar que esta solución es también solución del sistema original (1).

El proceso algorítmico para encontrar las soluciones a un sistema lineal que acabamos de ejemplificar, se denomina *algoritmo de eliminación gaussiana*, y lo formalizamos enseguida.

a. Algoritmo de eliminación gaussiana

Si observamos el proceso que llevamos a cabo para encontrar la solución del sistema de ecuaciones (1), caemos en la cuenta de que fue efectuado *sobre los coeficientes de las variables*. Esta observación nos lleva a introducir una "matriz" donde sólo aparecen los coeficientes del sistema de ecuaciones. Consideremos, de nuevo, el sistema de ecuaciones (1):

$$\begin{aligned} x+3y-z &= 1 \\ 3x-y+z &= 0 \\ x+y+z &= 2 \end{aligned} \tag{1}$$

A este sistema podemos asociarle una "matriz" de coeficientes que llamaremos la *"matriz aumentada"*. Esta matriz consiste en los coeficientes de las variables y en las constantes que se encuentran en el lado derecho de las ecuaciones, colocados en el mismo orden en que aparecen en el sistema. Por ejemplo, *la matriz aumentada* de este sistema (1) es

$$\begin{array}{c} \left[\begin{array}{rrr|r} 1 & 3 & -1 & 1 \\ 3 & -1 & 1 & 0 \\ 1 & 1 & 1 & 2 \end{array}\right] \end{array} \begin{array}{l} F_1 \\ F_2 \\ F_3 \end{array}$$

Aquí, las letras F_1, F_2, F_3 representan las filas 1, 2, 3 de la matriz, respectivamente. Esto nos permitirá indicar, con claridad, las operaciones que efectuaremos sobre cada fila.

Para encontrar la solución del sistema podemos llevar a cabo todo el proceso que efectuamos en la sección anterior, utilizando esta matriz. Se tienen tres *operaciones entre filas* que se pueden llevar a cabo:

1. Multiplicar todos los elementos de una fila por un escalar no nulo k.
2. Sumar o restar un múltiplo escalar de una fila a otra.
3. Intercambiar dos filas.

Las operaciones que acabamos de enunciar se conocen como *operaciones elementales entre filas* o, simplemente, *operaciones fila*.

En nuestro caso, sumemos la segunda fila con -3 veces la primera fila y el resultado lo reemplazamos por la segunda fila. Esto lo indicamos por $F_2 \longleftrightarrow F_2 - 3F_1$. Estas operaciones generan la siguiente matriz:

$$\left[\begin{array}{rrr|r} 1 & 3 & -1 & 1 \\ 0 & -10 & 4 & -3 \\ 1 & 1 & 1 & 2 \end{array}\right] \begin{array}{l} \\ F_2 \longleftrightarrow F_2 - 3F_1 \\ \\ \end{array}$$

Podemos ahora restar la primera fila de la tercera y reemplazarla por la tercera fila ($F_3 \longleftrightarrow F_3 - F_1$), para obtener la siguiente matriz:

$$\left[\begin{array}{rrr|r} 1 & 3 & -1 & 1 \\ 0 & -10 & 4 & -3 \\ 0 & -2 & 2 & 1 \end{array}\right] \begin{array}{l} \\ \\ F_3 \longleftrightarrow F_3 - F_1 \end{array}$$

Luego, multipliquemos por $-\frac{1}{10}$ la segunda fila de esta nueva matriz y la reemplazamos por la segunda fila ($F_2 \longleftrightarrow -\frac{1}{10}F_2$). El resultado es:

$$\left[\begin{array}{rrr|r} 1 & 3 & -1 & 1 \\ 0 & 1 & -\frac{2}{5} & \frac{3}{10} \\ 0 & -2 & 2 & 1 \end{array}\right] \begin{array}{l} \\ F_2 \longleftrightarrow -\frac{1}{10}F_2 \\ \\ \end{array}$$

Ahora multipliquemos por 2 la segunda fila y sumémosla a la tercera fila, y el resultado lo colocamos en reemplazo de la tercera fila ($F_3 \longleftrightarrow F_3 + 2F_2$).

Estas operaciones generan la siguiente matriz:

$$\begin{bmatrix} 1 & 3 & -1 & | & 1 \\ 0 & 1 & -\frac{2}{5} & | & \frac{3}{10} \\ 0 & 0 & \frac{6}{5} & | & \frac{8}{5} \end{bmatrix} F_3 \longleftrightarrow F_3 + 2F_2$$

Finalmente, multipliquemos la tercera fila por $\frac{5}{6}$ y el resultado lo colocamos como tercera fila ($F_3 \longleftrightarrow \frac{5}{6}F_3$). El resultado es:

$$\begin{bmatrix} 1 & 3 & -1 & | & 1 \\ 0 & 1 & -\frac{2}{5} & | & \frac{3}{10} \\ 0 & 0 & 1 & | & \frac{4}{3} \end{bmatrix} F_3 \longleftrightarrow \frac{5}{6}F_3$$

De esta forma podemos concluir que el sistema de ecuaciones lineales (1) es posible transformarlo, mediante operaciones fila, en el sistema

$$\begin{aligned} x + 3y - z \quad &= 1 \\ y - \frac{2}{5}z \quad &= \tfrac{3}{10} \\ z \quad &= \tfrac{4}{3} \end{aligned}$$

Notemos que en este punto ya podemos leer el valor de z en la última ecuación. Haciendo *sustitución hacia atrás*, podemos encontrar los valores que toman las demás variables. Así, $z = \frac{4}{3}$, $y = \frac{5}{6}$ y $x = -\frac{1}{6}$; y comprobamos que, efectivamente, es solución al sistema (1) original.

Ejemplo 1. (Otro ejemplo de solución por eliminación gaussiana) Consideremos ahora el siguiente sistema de ecuaciones:

$$\begin{aligned} 2x + 2y + 3z + 3w &= 4 \\ x + y + z + 2w &= 1 \\ 2x + 3y + 4z + 5w &= 2 \\ x + 2y + 3z + 10w &= 8 \end{aligned} \tag{6}$$

La matriz aumentada de este sistema es:

$$\begin{bmatrix} 2 & 2 & 3 & 3 & | & 4 \\ 1 & 1 & 1 & 2 & | & 1 \\ 2 & 3 & 4 & 5 & | & 2 \\ 1 & 2 & 3 & 10 & | & 8 \end{bmatrix} \begin{array}{l} F_1 \\ F_2 \\ F_3 \\ F_4 \end{array}$$

En un primer paso, en la matriz llevemos el coeficiente de la primera fila y de la primera columna, 2, a 1. Esto se puede efectuar de varias maneras: podríamos multiplicar toda la primera fila por $\frac{1}{2}$ o podríamos intercambiar las filas primera y segunda de la matriz. Observemos que cada uno de estos dos procedimientos corresponde a una operación elemental entre filas. En este ejercicio escogemos efectuar el cambio de las filas 1 y 2:

$$\begin{bmatrix} 1 & 1 & 1 & 2 & | & 1 \\ 2 & 2 & 3 & 3 & | & 4 \\ 2 & 3 & 4 & 5 & | & 2 \\ 1 & 2 & 3 & 10 & | & 8 \end{bmatrix} \begin{array}{l} F_1 \longleftrightarrow F_2 \\ F_2 \longleftrightarrow F_1 \\ \\ \\ \end{array}$$

Luego, restemos de la segunda fila y la tercera fila, dos veces la primera fila; y de la cuarta fila restemos la primera fila. Estas operaciones producen la siguiente matriz:

$$\begin{bmatrix} 1 & 1 & 1 & 2 & | & 1 \\ 0 & 0 & 1 & -1 & | & 2 \\ 0 & 1 & 2 & 1 & | & 0 \\ 0 & 1 & 2 & 8 & | & 7 \end{bmatrix} \begin{array}{l} \\ F_2 \longleftrightarrow F_2 - 2F_1 \\ F_3 \longleftrightarrow F_3 - 2F_1 \\ F_4 \longleftrightarrow F_4 - F_1 \end{array}$$

Ahora podemos ubicar un 1 en la casilla de la segunda fila y segunda columna, intercambiando las filas 2 y 3:

$$\begin{bmatrix} 1 & 1 & 1 & 2 & | & 1 \\ 0 & 1 & 2 & 1 & | & 0 \\ 0 & 0 & 1 & -1 & | & 2 \\ 0 & 1 & 2 & 8 & | & 7 \end{bmatrix} \begin{array}{l} \\ F_2 \longleftrightarrow F_3 \\ F_3 \longleftrightarrow F_2 \\ \\ \end{array}$$

Ya podemos eliminar las casillas de la cuarta fila con la segunda columna, y de la cuarta fila con la tercera columna, restando, de la cuarta fila, la segunda. Esto produce la siguiente matriz:

$$\begin{bmatrix} 1 & 1 & 1 & 2 & | & 1 \\ 0 & 1 & 2 & 1 & | & 0 \\ 0 & 0 & 1 & -1 & | & 2 \\ 0 & 0 & 0 & 7 & | & 7 \end{bmatrix} \begin{array}{l} \\ \\ \\ F_4 \longleftrightarrow F_4 - F_2 \end{array}$$

Finalmente, multipliquemos la última fila por $\frac{1}{7}$ para obtener que:

$$\begin{bmatrix} 1 & 1 & 1 & 2 & | & 1 \\ 0 & 1 & 2 & 1 & | & 0 \\ 0 & 0 & 1 & -1 & | & 2 \\ 0 & 0 & 0 & 1 & | & 1 \end{bmatrix} \begin{array}{l} \\ \\ \\ F_4 \longleftrightarrow \frac{1}{7}F_4 \end{array}$$

De esta forma, podemos concluir que el sistema de ecuaciones lineales resultante es:

$$\begin{aligned} x + y + z + 2w &= 1 \\ y + 2z + w &= 0 \\ z - w &= 2 \\ w &= 1 \end{aligned}$$

Por sustitución hacia atrás, obtenemos que $w = 1$, $z = 3$, $y = -7$ y $x = 3$; y confirmamos que también es solución al sistema (6) original. ▲

Pero debería aquí, sin embargo, surgir la pregunta sobre si las operaciones fila aplicadas a un sistema de ecuaciones lineales, afectan o no a sus soluciones. La respuesta la encontramos en el siguiente teorema general que revela la característica fundamental de los sistemas de ecuaciones lineales.

Teorema 1. ***(Operaciones fila y soluciones)***
Si se aplica cualquiera de las operaciones elementales entre filas a un sistema de ecuaciones lineales, el nuevo sistema que se obtiene posee exactamente las mismas soluciones del sistema original.

Demostración
Consideremos el sistema lineal de m ecuaciones con n incógnitas

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
\vdots \qquad \vdots \qquad \vdots \qquad \vdots \quad &\ \ \vdots \\
a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n &= b_k \\
\vdots \qquad \vdots \qquad \vdots \qquad \vdots \quad &\ \ \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{aligned} \tag{7}$$

y también los sistemas

$$i) \qquad \begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
\vdots \qquad \vdots \qquad \vdots \qquad \vdots \quad &\ \ \vdots \\
\alpha\, a_{k1}x_1 + \alpha\, a_{k2}x_2 + \cdots + \alpha\, a_{kn}x_n &= \alpha\, b_k \\
\vdots \qquad \vdots \qquad \vdots \qquad \vdots \quad &\ \ \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{aligned} \tag{8}$$

(En el anterior sistema, multiplicamos la k-ésima ecuación por el escalar $\alpha \neq 0$)

$$
ii) \qquad
\begin{array}{rr}
 & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\
 & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\
 & \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
\text{posición } k \longrightarrow & (\,\alpha\, a_{k1} + a_{l1}\,)x_1 + (\,\alpha\, a_{k2} + a_{l2}\,)x_2 + \cdots + (\,\alpha\, a_{kn} + a_{ln}\,)x_n = \alpha\, b_k + b_l \\
 & \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
 & a_{l1}x_1 + a_{l2}x_2 + \cdots + a_{ln}x_n = b_l \\
 & \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
 & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m
\end{array}
\tag{9}
$$

(sumar α veces la ecuación k más la ecuación l, colocada en la posición de la ecuación k)

$$
iii) \qquad
\begin{array}{rr}
 & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\
 & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\
 & \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
\text{posición } k \longrightarrow & a_{l1}x_1 + a_{l2}x_2 + \cdots + a_{ln}x_n = b_l \\
 & \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
\text{posición } l \longrightarrow & a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n = b_k \\
 & \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
 & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m
\end{array}
\tag{10}
$$

(intercambio entre las ecuaciones k y l).

Comparemos las soluciones de (7) y (8); de (7) y (9); y de (7) y (10):

a) Como $\alpha \neq 0$, las soluciones de (7) y (8) son exactamente las mismas.

b) De un lado, si $x^* = (\,x_1^*,\, x_2^*,\, \ldots,\, x_n^*\,)$ es solución de (7), entonces, puesto que $a_{k1}x_1^* + a_{k2}x_2^* + \cdots + a_{kn}x_n^* = b_k$ y $a_{l1}x_1^* + a_{l2}x_2^* + \cdots + a_{ln}x_n^* = b_l$, se tiene que

$$(\,\alpha\, a_{k1} + a_{l1}\,)x_1^* + (\,\alpha\, a_{k2} + a_{l2}\,)x_2^* + \cdots + (\,\alpha\, a_{kn} + a_{ln}\,)x_n^* = \alpha\, b_k + b_l;$$

luego x^* también es solución de (9). De otro lado, si x^* es solución de (9), entonces es solución de

$$(\,\alpha\, a_{k1} + a_{l1}\,)x_1 + (\,\alpha\, a_{k2} + a_{l2}\,)x_2 + \cdots + (\,\alpha\, a_{kn} + a_{ln}\,)x_n = \alpha\, b_k + b_l$$

y de

$$a_{l1}x_1 + a_{l2}x_2 + \cdots + a_{ln}x_n = b_l,$$

luego (restando ambas ecuaciones) es solución de

$$\alpha\, a_{k1}x_1 + \alpha\, a_{k2}x_2 + \cdots + \alpha\, a_{kn}x_n = \alpha\, b_k$$

y como $\alpha \neq 0$, x^* es solución de

$$a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n = b_k$$

y así x^* es solución de (7).

c) El intercambio de ecuaciones, obviamente, no afecta las soluciones del sistema (7). ■

b. Una visión geométrica

Como veremos en una lección más adelante, para un sistema de ecuaciones *lineales* sólo puede suceder que haya una única solución, no hayan soluciones o existan infinitas soluciones: *este hecho es una característica esencial de la linealidad*. Nunca encontraremos (como sí sucede en sistemas no-lineales) que el sistema tenga, por ejemplo, *dos* soluciones.

Podemos mostrar, *geométricamente*, situaciones en el plano en donde se indican con ejemplos los tres casos posibles que se pueden presentar al resolver un sistema de ecuaciones:

i) Consideremos inicialmente el sistema

$$\begin{aligned} 3x - 4y &= 4 \\ x + 2y &= 6 \end{aligned}$$

Este sistema tiene solución única, $\left(\frac{16}{5}, \frac{7}{5}\right)$, que se ve gráficamente donde las dos rectas se cruzan (ver figura 1).

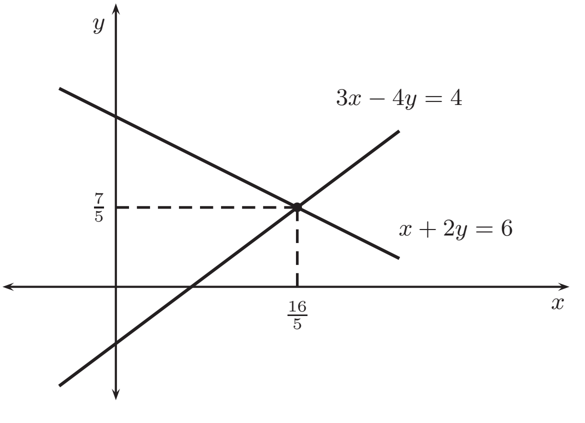


Figura 1. Solución única a un sistema lineal

ii) El sistema de ecuaciones

$$\begin{aligned} x+2y &= 6 \\ 3x+6y &= 4 \end{aligned}$$

representa el par de rectas paralelas que se muestra en la figura 2 y, por tanto, este sistema no tiene solución.

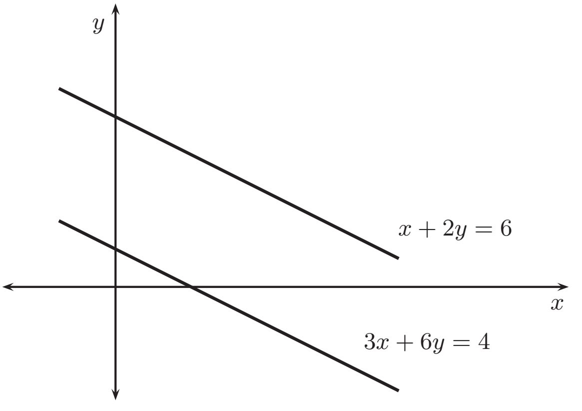


Figura 2. Sistema lineal sin soluciones

iii) Consideremos ahora el sistema de ecuaciones

$$\begin{aligned} x+2y &= 6 \\ 3x+6y &= 18 \end{aligned}$$

En este caso, ambas ecuaciones representan *exactamente la misma recta*, así que el sistema tiene un número infinito de soluciones.

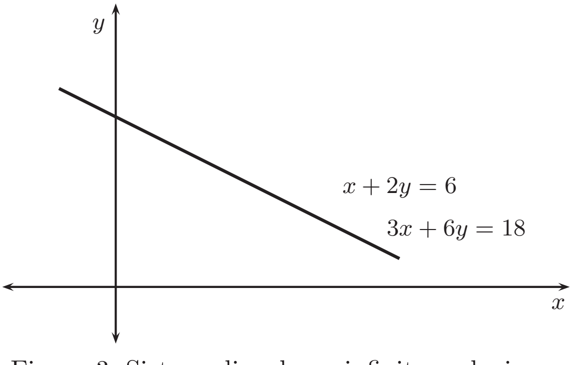


Figura 3. Sistema lineal con infinitas soluciones

Ejemplo 2.
Encontremos los valores de k tales que el sistema

$$\begin{aligned} x-y &= 5 \\ 3x+ky &= 7 \end{aligned}$$

a) Tenga sólo una solución.

b) No tenga solución.

c) Tenga infinitas soluciones.

Solución

La matriz aumentada del sistema es

$$\left[\begin{array}{cc|c} 1 & -1 & 5 \\ 3 & k & 7 \end{array}\right]$$

Sumando la segunda fila y -3 veces la primera se tiene la siguiente matriz:

$$\left[\begin{array}{cc|c} 1 & -1 & 5 \\ 0 & 3+k & -8 \end{array}\right] F_2 \leftrightarrow F_2 - 3F_1$$

Por tanto, no existe solución si $k = -3$, y la solución, dividiendo por $k+3$ en la matriz anterior, es única si $k \neq -3$. ¿Cuál es la solución del sistema en este último caso? Observemos que no existe un valor de k para el cual las soluciones sean infinitas. ▲

Ejemplo 3. (Solución a un sistema general)

Apliquemos la eliminación gaussiana para estudiar el sistema general

$$\begin{aligned} a_{11}x + a_{12}y &= b_1 \\ a_{21}x + a_{22}y &= b_2 \end{aligned}$$

con a_{11}, a_{12}, a_{21}, $a_{22} \in \mathbb{R} - \{0\}$. ¿Cuál es la condición sobre $a_{11}\,a_{22} - a_{12}a_{21}$ para que la solución sea única?

Solución

La matriz aumentada de este sistema general es

$$\left[\begin{array}{cc|c} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \end{array}\right]$$

Multiplicando por $\dfrac{1}{a_{11}}$ la primera fila obtenemos

$$\left[\begin{array}{cc|c} 1 & \frac{a_{12}}{a_{11}} & \frac{b_1}{a_{11}} \\ a_{21} & a_{22} & b_2 \end{array}\right] F_1 \leftrightarrow \frac{1}{a_{11}}F_1$$

Sumando la segunda fila y $-a_{21}$ veces la primera se tiene que

$$\left[\begin{array}{cc|c} 1 & \frac{a_{12}}{a_{11}} & \frac{b_1}{a_{11}} \\ 0 & -\frac{a_{12}a_{21}}{a_{11}} + a_{22} & -\frac{a_{21}b_1}{a_{11}} + b_2 \end{array}\right] F_2 \leftrightarrow F_2 - a_{21}F_1$$

Por lo tanto, la condición para que la solución de este sistema de ecuaciones lineales sea única es $-\dfrac{a_{12}a_{21}}{a_{11}} + a_{22} \neq 0$; es decir, $a_{11}\, a_{22} - a_{12}a_{21} \neq 0$. ¿Cuáles serían las soluciones x y y en este caso? ▲

También podemos mostrar *geométricamente*, pero ahora en situaciones en el espacio, los tres casos posibles de soluciones a un sistema de tres ecuaciones lineales con tres incógnitas. Imagine cada ecuación representando un plano en el espacio (Volumen 0 (Fundamentos)). Se tendrán, por ejemplo, tres planos paralelos (figura 4); dos planos yuxtapuestos y el otro paralelo (figura 5); los tres planos intersectándose en una recta (figura 6); los tres planos intersectándose en un punto (figura 8); etc. ¿Qué otras situaciones podrían suceder?

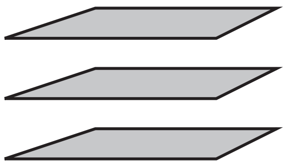


Figura 4. Tres planos paralelos; no hay solución.

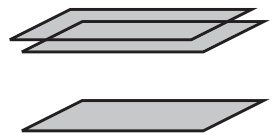


Figura 5. Dos planos yuxtapuestos, y el otro paralelo; no hay solución.

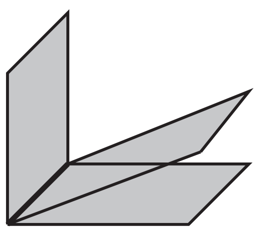


Figura 6. Tres planos que se intersectan en una recta; infinitas soluciones.

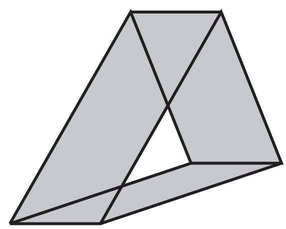


Figura 7. Tres planos se intersectan dos a dos; no hay solución.

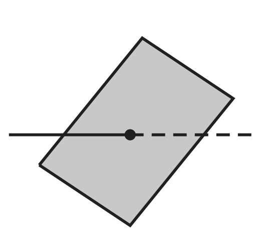


Figura 8. La recta es la intersección de dos planos. La intersección de esta recta con el plano que aparece en la figura da como resultado un sólo punto; esta es la solución al sistema.

Ejemplo 4. (Un sistema sin solución)
Resolvamos, si es posible,

$$\begin{aligned} x + 2y + 3z &= 4 \\ 4x + 5y + 6z &= 7 \\ 7x + 8y + 9z &= 12 \end{aligned} \tag{11}$$

Solución

Aplicando el método de eliminación gaussiana a (11) obtenemos:

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 4 & 5 & 6 & 7 \\ 7 & 8 & 9 & 12 \end{array}\right] \begin{array}{l} F_1 \\ F_2 \\ F_3 \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & -3 & -6 & -9 \\ 0 & -6 & -12 & -16 \end{array}\right] \begin{array}{l} \\ F_2 \leftrightarrow F_2 - 4F_1 \\ F_3 \leftrightarrow F_3 - 7F_1 \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & -3 & -6 & -9 \\ 0 & 0 & 0 & 2 \end{array}\right] \begin{array}{l} \\ \\ F_3 \leftrightarrow F_3 - 2F_2 \end{array}$$

Pero de esta última matriz ampliada encontramos entonces la ecuación $0 = 2$ que, obviamente, no puede cumplirse, independientemente de los valores de x, y y z. Como el último sistema es equivalente al primero, el sistema (11) *no tiene solución*.

Ejemplo 5. (Un sistema con infinitas soluciones)
Resolvamos el siguiente sistema de ecuaciones lineales:

$$\begin{aligned} x + 2y + 3z &= 0 \\ 4x + 5y + 6z &= 0 \\ 7x + 8y + 9z &= 0 \\ 10x + 11y + 12z &= 0 \end{aligned} \tag{12}$$

Solución

La matriz aumentada de este sistema es

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 0 \\ 10 & 11 & 12 & 0 \end{array}\right] \begin{array}{l} F_1 \\ F_2 \\ F_3 \\ F_4 \end{array}$$

Aplicando el método de eliminación gaussiana a (12) obtenemos:

$$\left[\begin{array}{rrr|r} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 0 \\ 0 & -6 & -12 & 0 \\ 0 & -9 & -18 & 0 \end{array}\right] \begin{array}{c} \\ F_2 \leftrightarrow -4F_1 + F_2 \\ F_3 \leftrightarrow -7F_1 + F_3 \\ F_4 \leftrightarrow -10F_1 + F_4 \end{array}$$

$$\left[\begin{array}{rrr|r} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & -6 & -12 & 0 \\ 0 & -9 & -18 & 0 \end{array}\right] \begin{array}{c} \\ F_2 \leftrightarrow -\dfrac{1}{3}F_2 \\ \\ \\ \end{array}$$

$$\left[\begin{array}{rrr|r} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{c} \\ \\ F_3 \leftrightarrow 6F_2 + F_3 \\ F_4 \leftrightarrow 9F_2 + F_4 \end{array}$$

De aquí se obtiene

$$\begin{aligned} x + 2y + 3z &= 0 \\ y + 2z &= 0 \end{aligned}$$

Y haciendo $z = t$, se llega a que

$$\begin{aligned} &y = -2t \\ &x = -2y - 3z = 4t - 3t = t \end{aligned}$$

y así se tienen *infinitas soluciones*: $x = t,\ y = -2t,\ z = t$; donde t varía sobre todos los números reales.

Nota 2. (Sobre el origen del método gaussiano)
El método de eliminación gaussiano fue utilizado por Gauss en un trabajo de 1809 en el que estudiaba la órbita de cierto asteroide (*Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium*). Utilizando observaciones tomadas desde 1803 hasta 1809, Gauss obtuvo un sistema de seis ecuaciones lineales con seis incógnitas, y allí mismo desarrolló un método sistemático para resolver este tipo de ecuaciones que es similar al método de eliminación que acabamos de estudiar.

Ejercicios 2

1) Resuelva y dibuje los siguientes sistemas de ecuaciones:

a) $$\begin{aligned} 4x + 6y &= 2 \\ 2x - 3y &= 5 \end{aligned}$$

b) $$\begin{aligned} 3x - y &= 9 \\ 5x + 3y &= -6 \end{aligned}$$

2) Resuelva, si es posible, los siguientes sistemas de ecuaciones:

a) $$\begin{aligned} x - 3y + z &= 5 \\ 2x + y - 2z &= -1 \\ x + 4y - 4z &= 0 \end{aligned}$$

b) $$\begin{aligned} x + 2y &= 4 \\ 3x + 6y &= 7 \\ 5x + 9y &= 12 \end{aligned}$$

c) $$\begin{aligned} 4x + 2y - 3z &= 1 \\ 6x + 3y - 5z &= 0 \\ x + y + 2z &= 9 \end{aligned}$$

d) $$\begin{aligned} x + y &= -1 \\ x + 2y - 2z &= 1 \\ -x - 2y + 3z &= 1 \end{aligned}$$

e) $$\begin{aligned} -2x + y + 7z &= 3 \\ 4x + 9y + 2z &= -6 \end{aligned}$$

f) $$\begin{aligned} x - y + 2z &= 7 \\ 3x + z &= 0 \\ 2x - y + z &= 4 \end{aligned}$$

3) Calcule los valores de k tales que el sistema

$$\begin{aligned} x + 2y + z &= 1 \\ x + 2ky + z &= -2 \\ 3x + y - z &= 7 \end{aligned}$$

a) Tenga sólo una solución.
b) No tenga solución.
c) Tenga infinitas soluciones.

4) En cierto número de tres dígitos, el segundo dígito es igual a la suma del primero y del tercero; la suma del segundo y tercer dígitos es 8, y, si el primero y el tercer dígito se intercambian, el número aumenta en 99. Encuentre el número.

5) Un individuo quiere invertir 6 millones de pesos entre un activo que le da un interés del 10 % y otro activo que le da un interés del 6 %. ¿Cuánto debe invertir el individuo en cada activo para que el rendimiento total sea de $\$440,000$?

6) Un inversionista quiere comprar acciones tipo A, tipo B y tipo C. Si compra tres acciones tipo A, dos acciones tipo B, y cinco tipo C, debe pagar 15 millones de pesos. Si compra cinco de cada tipo, debe pagar 28 millones. Y si compra 20 tipo C deberá pagar lo mismo que pagaría por tres acciones tipo A y dos tipo B. Muestre que el valor de las acciones es 2 millones para el tipo A, 3 millones para el tipo B, y $600,000 para el tipo C.

7) Un contratista tiene 30 obreros a su cargo, distribuidos en 3 categorías: A, B y C. El número de empleados en B es el doble que en A y C juntos. A los de la categoría A se les paga $16,000 al día; a los de la categoría B se les paga $20,000 al día; y a los de la categoría C se les paga $24,800 al día. La nómina diaria del contratista es de $586,400. ¿Cuántos empleados hay en cada categoría?

3. Contexto económico

a. Sobre el álgebra lineal en la teoría económica

Pensar en "términos lineales" como idealización de los distintos procesos económicos de una sociedad tiene una larga y respetable tradición en la historia de la teoría económica desde, por lo menos, François Quesnay [1694-1774] y su *Tableau Économique* de 1758, pasando por David Ricardo [1772-1823] y Karl Marx [1818-1883]. Aún así, ninguno de los economistas clásicos (a pesar de desarrollar discusiones y cálculos aritméticos o algebraicos que podrían asociarse con condiciones de linealidad) alcanzó ningún nivel formal de análisis en esta dirección. Más aún, a través del período de la economía matemática que va desde su fundador Augustin Louis Cournot [1801-1877] hasta los 1930's de John Maynard Keynes [1883-1946], la principal herramienta matemática de la teoría económica (con algunas excepciones) fue, de hecho, el *cálculo diferencial* de Newton y Leibniz y no los métodos del álgebra lineal.

Y esto era entendible, pues muchos de los problemas económicos eran planteados como problemas de máximos y mínimos, y el cálculo diferencial es la herramienta estándar a utilizar en tales casos, a pesar de que se sabe que existen problemas muy elementales de maximización que *no* son solubles mediante las técnicas diferenciales típicas. La *teoría económica lineal*, es decir, la teoría económica que utiliza el álgebra lineal, deriva mucho de su carácter del hecho de que es una aplicación de métodos del álgebra lineal no abarcados por el cálculo (desarrollados, básicamente, de los años treinta a los sesenta) a la economía. *El análisis insumo-producto, la teoría de juegos, y el análisis de actividades*, que describiremos adelante, fueron considerados ejemplos de una nueva perspectiva para una parte central de la teoría económica convencional, a la vez que sus interesantes aplicaciones la mostraban un poco más ligera que las técnicas del cálculo diferencial.

i) El primer modelo "lineal" (formal) importante en la historia del pensamiento económico es el hoy conocido como *modelo Walras-Cassel* de 1918 presentado por Gustave Cassel [1886-1945] en alemán en *Theoretische Sozialökomie*, y después traducido al inglés bajo el título *The Theory of Social Economy* en 1932. Para Léon Walras [1834–1910] en la primera y segunda edición de sus *Élements d'Economie Politique Pure (1874-77), la economía era un sistema en equilibrio en el que la producción se ajustaba a la demanda a través del mecanismo de precios*, y para esto había desarrollado complicados (para la época) sistemas de ecuaciones que, al resolverse simultáneamente, arrojaban el sistema de precios de bienes y factores (insumos) de la economía que la colocaba en ese equilibrio. Posteriormente, Cassel, utilizando técnicas propias de lo que hoy conocemos como álgebra lineal, encontró en 1918 que el inmenso aparato de Walras, aún en sencillos y significativos casos, *simplemente no funcionaba* y trató de remediarlo. Sin embargo, Cassel nunca supo presentar una demostración matemática de existencia de solución al sistema walrasiano aunque su trabajo destacó la necesidad de recurrir, en lugar de ecuaciones lineales de la forma $ax + by = c$, a desigualdades lineales de la forma $ax + by \leq c$.

Pero en aquella época de principios del siglo XX, ya los métodos matemáticos lineales eran suficientes para tratar con el cálculo de equilibrios en sistemas que están restringidos no sólo por igualdades sino por desigualdades lineales. Uno de los intentos más notables por probar la existencia de un equilibrio para el sistema Walras-Cassel extendido fue el del matemático Abraham Wald en 1935 (*Ergebnisse Eines Mathematischer Kolloquiums*). Sin embargo, se mostró que los argumentos de Wald no eran del todo satisfactorios pues imponían fuertes hipótesis sobre el modelo y esto era contrario al espíritu del modelo de Cassel. En cualquier caso, el sistema lineal extendido de Wald fue susceptible de cierta manipulación matemática y el modelo de equilibrio general lineal tendría posteriormente su formulación definitiva y generalización en los trabajos de Koopmans (1951)[1], McKenzie (1954)[2] y Arrow y Debreu (1954)[3]. Sobre el modelo Walras-Cassel discutiremos en la lección 4.

ii) El segundo modelo lineal importante en la historia del pensamiento económico fue el *análisis insumo-producto* del premio Nobel en economía de 1973, Wassily W. Leontief [1906-1999]. En *Quantitative Input and Output Relations in the Economic System of the United States* de 1936, Leontief mostraba su análisis insumo-producto como una consecuencia práctica de la teoría walrasiana que postula la interdependencia general de las variables económicas. Y aunque la aplicación es muy simple, la construcción de una *tabla* (matriz) *insumo-producto* puede implicar muchos problemas y resulta en extremo laboriosa. Las primeras tablas insumo-producto se construyeron en los años treinta para la economía de los Estados Unidos. Sobre esto regresaremos en la lección 5.

iii) Muy importante también para la economía lineal y de profunda influencia en el desarrollo del pensamiento económico ha sido el famoso artículo

[1] Koopmans, Tjalling C. (1951), *Activity Analysis of Production and Allocation*, New York: Wiley.

[2] McKenzie, Lionel (1954), On Equilibrium in Graham's Model of World Trade and Other Competitive Systems, *Econometrica*, vol. 22, 147-61.

[3] Arrow, K. y G. Debreu (1954), Existence of an Equilibrium for a Competitive Economy, *Econometrica*, vol. 22, 265-290.

de John von Neumann [1903-1957] "A Model of General Economic Equilibrium" de 1946 (que fuera presentado en 1932). En él, von Neumann desarrolla un modelo dinámico de una economía en expansión construido como una sucesión de *períodos interconectados*, que a su vez los considera (separadamente) como sistemas *estáticos*. La forma en que estudia estos períodos estáticos es, sustancialmente, a la manera del sistema Walras-Cassel extendido de Wald, pero que hábilmente los relaciona con un modelo similar al de Leontief (aunque no inspirado en éste último). El trabajo de von Neumann ha sido, sin duda, el padre de muchos de los modelos de crecimiento económico que hoy conocemos. Esta discusión la presentamos en la lección 6.

iv) Desde una orilla de aproximación diferente, Piero Sraffa [1898–1983] publica en 1960 *Production of Commodities by Means of Commodities* en el cual mostraba cómo era posible encontrar, para cualquier economía, una *medida de valor*, un problema que ya antes había preocupado a todos los economistas clásicos y, en particular, a Ricardo. Esta mercancía, que se conoce como *mercancía estándar (o patrón)*, permite entonces medir *precios relativos que no cambian* a menos que la tecnología que la produce, cambie. En el esquema de Sraffa la demanda no entra en el sistema y los precios están determinados por el costo de producción. Conclusiones como éstas han permitido exploraciones alternas a la teoría del equilibrio walrasiano; en particular, en la teoría del valor, en la teoría de la distribución del ingreso, y en la teoría del crecimiento económico. Sobre el modelo de Sraffa discutiremos hacia el final en la lección 7.

v) En 1936, John Maynard Keynes publicaba su más importante obra con la que culminaría su trabajo académico. *The General Theory of Employment, Interest and Money* rompe la tradición microeconómica que venía con Walras, William Jevons y Carl Menger (fortalecida por Vilfredo Pareto, Francis Edgeworth, Alfred Marshall y Arthur Pigou) sobre el comportamiento de los mercados específicos, sobre la asignación de recursos entre ellos, y sobre el ajuste rápido de precios buscando el equilibrio. Keynes, por su parte, estaba más interesado en el estudio de los eventos que afectan a *toda la economía* (inflación, deflación, auges, recesiones) y cuyo desarrollo conduciría a lo que hoy conocemos como *macroeconomía.* Aunque amplia, rica y profunda, en 1937 John R. Hicks [1904-1989] (premio Nobel en economía en 1972), propondría en *Mr. Keynes and the Classics: A Suggested Interpretation* una exposición matemática simplificada de la *Teoría General* de Keynes: el famoso modelo IS-LM. Fundamentalmente, el modelo trata de explicar la interacción entre los mercados reales y monetarios. Del mercado real extrae el nivel de ingresos (Y), y del mercado monetario extrae la tasa de interés (i), y cada una de estas variables, a su vez, afectan elementos en el otro mercado. Hicks, con "instintos walrasianos", cree entonces que la visión de Keynes coincide con la solución a un sistema *simultáneo* de ecuaciones: una para el mercado real, otra para el monetario y un equilibrio en el que ambas ecuaciones coinciden. Sobre la versión *lineal* de este modelo discutiremos en la lección 8.

vi) Después de los primeros desarrollos en economía lineal, estos se dispersaron en múltiples ramas, pero principalmente hacia la fundamentación de la teoría del equilibrio general walrasiano, como lo demuestra el soberbio análisis de equilibrio con herramientas lineales, del premio Nobel en economía de 1975, Tjalling C. Koopmans [1910-1985], quien en 1951 desarrollara lo que hoy conocemos en economía como *análisis de actividades*. El origen histórico de este importante trabajo se encuentra en las discusiones surgidas en los años treinta sobre las posibles generalizaciones del sistema walrasiano. Neisser (1932)[4] y von Stackelberg (1933)[5] abrirían debates con respecto a la existencia y unicidad de las soluciones del sistema walrasiano de Cassel, con particular referencia al problema de la posibilidad de precios negativos o nulos. Schlesinger (1933-4)[6] y Wald (1935, 1936, 1942)[7, 8, 9]probaron la existencia y unicidad de la solución a un sistema que expresaba este problema. Von Neumann (1932, 1945)[10] generalizó el modelo de Wald en varias direcciones como ya hemos mencionado. Pero la dirección más importante para Koopmans fue, quizás, resaltar una observación que von Neumann hacía al final del artículo de 1932: la solución al modelo era *eficiente* en el sentido de que la economía alcanzaba una tasa máxima de expansión de la producción, compatible con la tecnología dada de la producción y con el consumo.

---

4 Neisser, H. (1932), *Lohnhöhe und Beschäftigungsgrad im Marktgleichgewicht,* Weltwirtschaftliches Archiv.

5 Von Stackelberg, H. (1933), *Zwei Kritische Bemerkungen zur Preistheorie Gustav Cassels,* Zeitschrift für Nationalökonomie.

6 Schlesinger, K. (1933-4), Über die Produktionsgleichungen der Ökonomischen Wertlehre I, *Ergebnisse eines Mathematischen Kolloquiums*, vol. 6, 12-20.

7 Wald, A. (1935), Über die Eindeutige Positive Losbarkeit der Neuen Produktionsgleichungen, *Ergebnisse Eines Mathematischen Kolloquiums*, vol. 6, 12-18.

8 Wald, A. (1936), Über die Produktionsgleichungen der Ökonomischen Wetlehre II, *Ergebnisse eines Mathematischen Kolloquiums*, vol. 7, 1-6.

9 Wald, A. (1942), Über einige Gleichungssysteme der Mathematischen Ökonomie, *Zeitschrift für Nationalökonomie*, vol. 7, 637-670. Traducido como On Some Systems of Equations of Mathematical Economics (1951), *Econometrica*, vol. 19, 368-403.

10 Von Neumann, J. (1945), A Model of General Economic Equilibrium, *Review of Economic Studies*, vol. 13, 1-9.

Y es que la fuente del *bienestar económico* también inspiraría a Koopmans. El concepto de A. Bergson (1938)[11] de una *función de bienestar económico*, la *función de valor social* de Lange (1942)[12] y el *principio débil de bienestar* de Pareto (1906)[13], estaban en el centro de la discusión, pues el problema sobre si la noción de precio como "mensajero de información" que, circulando entre los distintos centros de decisión de una economía, hacía posible una distribución eficiente de los recursos, estaba en el corazón de las posibilidades de una economía socialista, y también de la de muchos sectores de las economías capitalistas o mixtas en donde los mercados no actúan.

Adicionalmente a lo anterior, otra fuente de inspiración para el análisis de actividades de Koopmans fue el trabajo sobre relaciones interindustriales iniciado, desarrollado y estimulado por Leontief. Uno de los propósitos de Koopmans (al igual que el de Leontief) fue el de que su modelo pudiera servir de base empírica para la estimación numérica de los efectos, sobre los niveles de actividad en las industrias individuales, de los cambios en la composición de la demanda final de la industrias.

Otra vertiente para Koopmans fue el estudio de ciertos trabajos igualmente prácticos (pero menos agregados) sobre asignación mediante modelos de *programación lineal* que surgieron, particularmente, en la organización de estrategias en la Segunda Guerra Mundial (1938–1945). El más completo trabajo en esta categoría fue el desarrollo de los modelos de programación lineal de George B. Dantzig y otros miembros de la Fuerza Aérea de los Estados Unidos. Estos modelos trataban aspectos dinámicos sobre cómo desarrollar actividades interdependientes de una organización grande, de tal manera que se alcanzara la mejor combinación hacia el logro de un objetivo establecido. La estructura conceptual era clara: *estudiar comparativamente ciertos objetivos alcanzables teniendo ciertos medios escasos al alcance.* Sobre el modelo de análisis (lineal) de actividades de Koopmans discutiremos en la lección 8.

vii) Sin embargo, una notable excepción a esta última línea fue la *teoría de juegos* del mismo von Neumann y del economista austríaco Oskar Morgenstern [1902–1976]. Los métodos del análisis lineal hasta los 1940's estaban atados, de alguna forma, a la hipótesis de *rendimientos constantes a escala* (Volumenn 0 (Fundamentos)) que se asocia a la ausencia

[11] Bergson, A. (1938), A Reformulation of Certain Aspects of Welfare Economics, *Quarterly Journal of Economics*, vol. 52, 310-334.

[12] Lange, O. (1942), The Foundations of Welfare Economics, *Econometrica*, vol. 10, 215-228.

[13] Pareto, V. (1906P), *Manual of Political Economy* (1971), traducción de la edición de 1927, New York: Augustus M. Kelley.

de economías de escala. Hasta aquella época, la economía lineal parecía haber mostrado que siempre que la hipótesis de rendimientos constantes a escala estuviera presente, también lo estaría la del mecanismo de precios. Pero la economía no es sólo una "relación entre fines y medios escasos que tienen usos alternativos" como decía el influyente economista inglés Lionel Robbins en *An Essay on the Nature and Significance of Economic Sciences* de 1935. Tiene que ver, en particular, con el "comportamiento humano" y con las "lógicas de elección". En *Theory of Games and Economic Behavior* (1944) de von Neumann y Morgenstern, aparece, por primera vez en el pensamiento económico formalizado, un análisis del comportamiento humano *cuando los fines son opuestos*. Es el primer esfuerzo formal por modelar *interacciones*, diferentes del indirecto mercado walrasiano. Los autores presentan aquí los primeros modelos de interacciones en donde el objetivo es encontrar sus "predicciones" (o equilibrios) del juego mediante ciertas "hipótesis de racionalidad" sobre los jugadores. Las técnicas lineales específicas sobre los modelos de la teoría de juegos de von Neumann y Morgenstern las estudiaremos en la lección 8.

Ejercicios complementarios

1. Resuelva, si es posible, los siguientes sistemas de ecuaciones lineales:

a) $$\begin{aligned} \frac{x}{3} + 5y &= 3 \\ 5x + 3y &= 1.65 \end{aligned}$$

b) $$\begin{aligned} 2(\,2x + 3y\,) &= 3(\,2x - 3y\,) + 10 \\ 4x - 3y &= 4(\,6y - 2x\,) + 3 \end{aligned}$$

c) $$\begin{aligned} ax + by &= c \\ px &= qy \end{aligned}$$

d) $$\begin{aligned} ax + by &= a^2 + 2a + b^2 \\ bx + ay &= a^2 + 2b + b^2 \end{aligned}$$

e) $$\begin{aligned} \frac{x-y}{4} - \frac{x+2y-5}{6} &= \frac{y-3}{4} - \frac{y+2x-5}{6} \\ 5x - 2y + 6 &= 0 \end{aligned}$$

2. Encuentre valores para a, b, c, d, e, f tales que el sistema

$$\frac{x}{a} + \frac{y}{b} = \frac{1}{c} \qquad\qquad \frac{x}{d} + \frac{y}{e} = \frac{1}{f}$$

a) Tenga solución única. b) Tenga infinitas soluciones.
c) No tenga ninguna solución.

3. a) Resuelva el sistema

$$\frac{2}{x} + \frac{5}{3y} = 1 \qquad\qquad \frac{9}{x} + \frac{10}{y} = 5$$

utilizando alguna sustitución que lo convierta en un sistema lineal. [Indicación: haga $\bar{x} \equiv \frac{1}{x}$, $\bar{y} \equiv \frac{1}{y}$]

b) Lo mismo, si lo considera necesario, para los sistemas siguientes:

a) $3x + \dfrac{y}{x} = 6; \quad 7x - \dfrac{2y}{x} = 1$

[Indicación: haga $\overline{y} = \frac{y}{x}$ y resuelva para x y $\overline{y}$].

b) $\dfrac{1}{2x} + \dfrac{1}{3y} = 0; \quad \dfrac{3}{x} + \dfrac{1}{y} + 3 = 0$

c) $x(\,x - y\,)(\,x + y\,) = 0; \quad x + 2y - 5 = 0$

d) $y^2 = (\,x - 1\,)^2; \quad 2x + 3y - 7 = 0$

e) $\dfrac{y}{x} = \dfrac{2(\,3 - y\,)}{x} + \dfrac{3}{2}; \quad \dfrac{y+3}{x} = \dfrac{3y-5}{x} + 1$

4. Resuelva los siguientes sistemas reduciéndolos (cuando sea necesario) a sistemas lineales:

a) $$\begin{aligned} x-y&=3\\ y-z&=-5\\ z-x&=2 \end{aligned}$$

b) $$\begin{aligned} 3x-8y+7z&=10\\ 2x+5y-3z&=12\\ 16x+9y-2&=\frac{3}{2} \end{aligned}$$

c) $$\begin{aligned} \frac{1}{x}+\frac{1}{y}-\frac{6}{z}&=9\\ \frac{1}{x}-\frac{1}{y}+\frac{4}{z}&=5\\ \frac{3}{y}-\frac{2}{x}-\frac{1}{z}&=4 \end{aligned}$$

d) $$\begin{aligned} y+z+u&=4\\ x+z+u&=3\\ x+y+u&=1\\ x+y+z&=10 \end{aligned}$$

e) $$\begin{aligned} 3x-5&=2(x-2)\\ (x+1)(y-1)&=(x+2)(y-2)+5\\ 2x+3y+2&=1 \end{aligned}$$

f) $$\begin{aligned} 3x^2+4y^2&=5\\ 5x^2-y^2&=2 \end{aligned}$$

5. Considere el siguiente sistema de ecuaciones lineales:

$$\begin{aligned} x-y+2z&=7\\ ax+z&=0\\ 2x-y+z&=b \end{aligned}$$

donde a y b son parámetros. Responda las siguientes preguntas utilizando el método gaussiano:

a) ¿Para qué valores de a y b existe más de una solución?

b) ¿Para qué valores de a y b no existe solución?

c) ¿Para qué valores de a y b existe una única solución? ¿Cuál es esta solución?

6. Encuentre tres números cuya suma sea 20 y que:
El primero, más dos veces el segundo, más tres veces el tercero, sea igual a 44; y que dos veces la suma del primero y el segundo, menos cuatro veces el tercero sea igual a -14.

7. Habíamos afirmado antes que alrededor del año 300 a. C., los babilonios estudiaban problemas que conducían a sistemas de ecuaciones lineales, y algunos de ellos se han preservado en tabletas de barro. Una de estas tabletas dice, en términos de hoy, lo siguiente: *Hay dos campos cuya área total es* $1,800$ *yardas cuadradas. Uno de ellos produce grano a una tasa*

de $\frac{2}{3}$ de bushel por yarda cuadrada, mientras que el otro produce grano a una tasa de $\frac{1}{2}$ bushel por yarda cuadrada. Si el campo total es de $1,100$ *bushels ¿cuál es la medida de cada campo?* Resuelva este problema de los babilonios.

8. Dos vasijas A y B, contienen mezclas de alcohol y agua. Una mezcla de tres partes de A y dos partes de B contendrá 40 % de alcohol; y una mezcla de una parte de A y dos partes de B contendrá 32 % de alcohol. Pruebe que los porcentajes de alcohol en A y B son 52 % y 22 %, respectivamente.

9. Una caja contiene 120 paquetes. Unos paquetes pesan media libra cada uno y los restantes un tercio de libra cada uno. Si el peso total del contenido de la caja es de 48 libras ¿cuántos paquetes de cada tipo hay?

10. Cuatro personas A, B, C, D se dividen $\$1'300,000$ de tal forma que B recibe $\frac{2}{3}$ de lo que recibe A; C recibe $\frac{2}{3}$ de lo que recibe B; y, D, $\frac{2}{3}$ de lo que recibe C. ¿Cuánto recibe cada uno?

11. Un granjero tiene 500 m^2 de terreno destinados al cultivo de maíz y trigo. El costo respectivo de los cultivos (incluyendo semillas y mano de obra) es de \$126,000 y \$90,000 por metro cuadrado. Si dispone de \$55'800,000 para realizar este cultivo y desea utilizar toda la tierra destinada a éstos, ¿cuántos metros cuadrados debe plantar de cada cultivo gastando todo el presupuesto?

12. ¿Será que un sistema (no necesariamente lineal) de m ecuaciones con m incógnitas tiene al menos una solución? Explique con ejemplos concretos.

13. Determine la solución de *equilibrio walrasiano* $(D_1 = S_1; D_2 = S_2)$ del modelo de mercado lineal de dos mercancías

$$\begin{aligned} D_1 &= 40 - 2P_1 + P_2; \qquad & S_1 &= 4P_1 - P_2 + 4 \\ D_2 &= 5P_1 - 2P_2 + 17; \qquad & S_2 &= 3P_2 - 4 \end{aligned}$$

    donde D, S y P significan *demanda*, *oferta* y *precio*, respectivamente; y 1 y 2 indican las mercancías. Justifique los signos positivos y negativos en cada una de las cuatro ecuaciones. Represente la solución gráficamente.

14. Una economía descrita mediante el modelo "keynesiano" IS-LM puede aparecer así:

$$i = \frac{1}{\beta}\,[F - (1-c)Y]$$

$$i = \frac{1}{h}\left[kY - \frac{M}{P}\right]$$

donde i = tasa de interés, F = "índice" de política fiscal (gasto), Y = ingreso agregado, M = saldos monetarios nominales, P = nivel de precios, y $\beta > 0$, $h > 0$, $k > 0$, $0 < c < 1$ son constantes. Si $\beta = 0.5$, $h = 0.3$, $k = 0.2$, $c = 0.9$, $\frac{M}{P} = 300$ y $F = 800$, calcule, si existen, los valores correspondientes de i y Y para que esta economía esté en "equilibrio". Represente la solución gráficamente.

Lección 2

Matrices y determinantes

Introducción

Sin duda, entre todas las funciones de una sola variable, las más simples son las *funciones lineales* $y = ax + b$, cuyas gráficas son también las curvas más simples: son *líneas rectas*. Una función lineal de una variable es tan sencilla en sus propiedades que no requiere de un estudio adicional al realizado en el Volumen 0 (Fundamentos). Sin embargo, el escenario se complica cuando estudiamos las funciones lineales en varias variables

$$y = a_1x_1 + a_2x_2 + \cdots + a_nx_n + b$$

y más aún cuando pasamos de una sola función en varias variables a un sistema de estas funciones:

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + b_1 \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + b_2 \\ \vdots \quad & \quad \vdots \qquad\quad \vdots \qquad\quad \vdots \qquad\quad \vdots \qquad \vdots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n + b_m \end{aligned}$$

El estudio de las funciones lineales y, especialmente, de los sistemas de ecuaciones lineales de la forma ya conocida

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad \vdots \quad & \quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m, \end{aligned}$$

constituyen el objeto inicial (y central) de estudio de lo que se ha dado en llamar *álgebra lineal*, y herramientas simplificadoras y fundamentales para este propósito son las *matrices* y los *determinantes*.

1. La noción de matriz

El sistema de ecuaciones lineales dado

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array} \tag{1}$$

se puede describir mediante su sistema de coeficientes ya que el *sistema depende únicamente de estos números y no de las letras con las que representemos las variables.* Por ejemplo, el sistema

$$\begin{aligned} 2x+3y-z&=5 \\ 3x+y+z&=1 \\ x-y+3z&=2 \end{aligned} \qquad \text{es equivalente al sistema} \qquad \begin{aligned} 2a+3b-c&=5 \\ 3a+b+c&=1 \\ a-b+3c&=2 \end{aligned}$$

El conjunto de coeficientes del sistema lineal (1) se acostumbra a reunir en dos ordenamientos rectangulares de la forma

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad \text{y} \qquad \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

y éste es el origen del concepto de matriz sobre el cual nos concentraremos durante esta lección, para posteriormente hacer un uso más amplio de él.

Definición 1. (Matriz (Silvester (1850), Cayley (1858)))
Una *matriz* (real) es un arreglo de números (reales) de la forma

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

donde a_{ij} se llamará la *entrada* ubicada en la i-ésima *fila* y la j-ésima *columna.* Para esta matriz, utilizaremos también la notación $A = [\,a_{ij}\,]_{m\times n}$, donde m es el número de filas y n es el número de columnas. El tamaño de una matriz de m filas y n columnas lo indicamos por $m \times n$. Si el número de filas coincide con el número de columnas, diremos que la matriz es *cuadrada.*

Nota 1.
La primera vez que se utilizó el término *matriz* fue en 1850 por el matemático J. J. Sylvester [1814–1897]. Éste definió una matriz como un "ordenamiento oblongo de términos". Sin embargo, Arthur Cayley [1821–1895], en 1858, publicó *Memoir on the Theory of Matrices* donde aparecía la *definición abstracta* de una matriz.

Ejemplo 1.

a) $\begin{bmatrix} -1 & 2 & 6 & 4 \\ 2 & 7 & 5 & 3 \\ 4 & 3 & 0 & 8 \end{bmatrix}$ es una matriz con 3 filas y 4 columnas.

b) $\begin{bmatrix} -3 & 3 & 0 & 1 \\ 8 & 2 & 8 & 9 \\ -5 & 4 & 6 & -4 \\ -1 & 7 & 5 & 3 \\ 0 & 1 & 2 & 6 \end{bmatrix}$ es una matriz con 5 filas y 4 columnas.

c) $\begin{bmatrix} 1 & 0 & 0 & -3 & 4 \\ 7 & 9 & 5 & 14 & 8 \end{bmatrix}$ es una matriz con 2 filas y 5 columnas.

Nota 2.
Las matrices de una única entrada, $[\,a\,]$, se identifican, normalmente, con el número a.

Definición 2. (Igualdad de matrices)
Diremos que dos matrices $A = [\,a_{ij}\,]_{m\times n}$ y $B = [\,b_{ij}\,]_{m\times n}$ son *iguales* si, y sólo si, sus respectivas entradas son iguales; es decir, si, y sólo si,

$$a_{ij} = b_{ij} \qquad \text{para} \quad i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n$$

Ejemplo 2.
Si

$$\begin{bmatrix} a+b & 2c-d \\ c+2d & a-b \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}$$

¿cuáles son $a,\ b,\ c,\ d$?

Solución

De acuerdo con la definición 2, los valores de $a,\ b,\ c,\ d$ están determinados por $a+b=1,\ a-b=2,\ 2c-d=3,\ c+2d=-1$. Por tanto, $a=\frac{3}{2},\ b=-\frac{1}{2}$, $c=1,\ d=-1$.

2. Tipos de matrices

En lo que sigue indicamos algunos de los tipos de matrices (y, por consiguiente, de sistemas de ecuaciones lineales) que aparecerán muy comúnmente en nuestro estudio.

Definición 3. (Matriz nula)
Una *matriz nula* o *matriz cero* es una matriz que tiene *todas* sus entradas iguales a cero. Se notará por $0_{m\times n}$ o, simplemente, 0, si esto no causa confusión.

Ejemplo 3.
Las siguientes son matrices nulas:

$$0_{2\times 2}=\begin{bmatrix}0&0\\0&0\end{bmatrix},\qquad 0_{2\times 3}=\begin{bmatrix}0&0&0\\0&0&0\end{bmatrix},\qquad 0_{3\times 3}=\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}$$

Definición 4. (Matriz idéntica)
Una matriz cuadrada que tiene en cada una de las entradas de su diagonal principal un 1 (uno) y en las demás entradas un 0 (cero) se llama *matriz idéntica* (o matriz identidad). La matriz idéntica de tamaño n la denotaremos por I_n. Observemos que podemos escribir I_n de la forma $[\,\delta_{ij}\,]_{n\times n}$, donde

$$\delta_{ij}=\begin{cases}1 & \text{si}\quad i=j\\ 0 & \text{si}\quad i\neq j\end{cases}$$

Ejemplo 4.
Las matrices idénticas I_2 y I_3 son

$$I_2=\begin{bmatrix}1&0\\0&1\end{bmatrix},\qquad I_3=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$$

Definición 5. (Matriz diagonal)
Una matriz cuadrada $A=[\,a_{ij}\,]_{n\times n}$ se llama *matriz diagonal* si todas las entradas que están por fuera de la diagonal principal son nulas; es decir, si $a_{ij}=0$ para $i\neq j$.

Ejemplo 5.
Las siguientes son matrices diagonales:

$$\begin{bmatrix}8&0\\0&9\end{bmatrix},\qquad \begin{bmatrix}1&0&0\\0&2&0\\0&0&3\end{bmatrix}$$

En particular, las matrices idénticas I_n son diagonales.

Definición 6. (Matriz triangular)
Una matriz cuadrada $A = [a_{ij}]_{n\times n}$ es una *matriz triangular superior* si todas las entradas que están por debajo de la diagonal principal son nulas; es decir, si $a_{ij} = 0$ para $j < i$. Análogamente, A es una *matriz triangular inferior* si todas las entradas que están por encima de la diagonal principal son nulas; es decir, si $a_{ij} = 0$ para $j > i$.

Claramente, una matriz es, simultáneamente, triangular superior e inferior si, y sólo si, es una matriz diagonal.

Ejemplo 6.

a) Las siguientes matrices son triangulares superiores:

$$\begin{bmatrix} 8 & 3 \\ 0 & 9 \end{bmatrix}, \qquad \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{bmatrix}$$

b) Las siguientes matrices son triangulares inferiores:

$$\begin{bmatrix} 8 & 0 \\ 3 & 9 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ -1 & 2 & 3 \end{bmatrix}$$

Ejercicios 2

1) Clasifique (mediante una tabla) las siguientes matrices de acuerdo a si son (o no) nulas, idénticas, diagonales, triangulares superiores y triangulares inferiores:

a) $\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$ b) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

c) $\begin{bmatrix} 3 & 2 & 1 \\ 0 & 4 & 0 \\ 0 & 0 & -5 \end{bmatrix}$ d) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

e) $\begin{bmatrix} 4 & 8 & 6 \\ 0 & 4 & -2 \\ 0 & 3 & 5 \end{bmatrix}$ f) $\begin{bmatrix} -3 & 0 & 0 \\ -15 & -5 & 0 \\ -12 & 1 & -10 \end{bmatrix}$

g) $\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$ h) $\begin{bmatrix} -10 & -3 & -2 \\ 0 & 0 & 8 \end{bmatrix}$

i) $\begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ 4 & 2 & 0 & 0 & 0 \\ -8 & 7 & -1 & 0 & 0 \\ 4 & -3 & -2 & -5 & 0 \\ 0 & 1 & 0 & 2 & -6 \end{bmatrix}$ j) $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

k) $\begin{bmatrix} -5 & -3 & -4 \\ 0 & 0 & -2 \\ 1 & 0 & 0 \end{bmatrix}$ l) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

3. Álgebra de matrices

En 1858, Arthur Cayley [1789-1857], en su ya mencionado *Memoir on the Theory of Matrices*, presentó también el álgebra de matrices al definir la suma, la multiplicación, la multiplicación por un escalar, e incluso la inversa (de la que hablaremos en la lección 3). Las tres operaciones algebraicas básicas que definió Cayley sobre las matrices, fueron: la suma de matrices, la multiplicación de matrices por escalar, y la multiplicación de matrices.

a. Suma de matrices

Supongamos que tenemos *dos* sistemas de ecuaciones lineales:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & y_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & y_2 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & y_m \end{array}$$

y

$$\begin{array}{ccccccccc} b_{11}x_1 & + & b_{12}x_2 & + & \cdots & + & b_{1n}x_n & = & z_1 \\ b_{21}x_1 & + & b_{22}x_2 & + & \cdots & + & b_{2n}x_n & = & z_2 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ b_{m1}x_1 & + & b_{m2}x_2 & + & \cdots & + & b_{mn}x_n & = & z_m \end{array}$$

Si sumamos la primera ecuación del primer sistema con la primera ecuación del segundo sistema; la segunda del primer sistema con la segunda del segundo sistema, etc., obtenemos que

$$\begin{array}{ccccccc}
(a_{11}+b_{11})x_1 &+& (a_{12}+b_{12})x_2 &+\cdots+& (a_{1n}+b_{1n})x_n &=& y_1+z_1\\
(a_{21}+b_{21})x_1 &+& (a_{22}+b_{22})x_2 &+\cdots+& (a_{2n}+b_{2n})x_n &=& y_2+z_2\\
\vdots & & \vdots & & \vdots & & \vdots\\
(a_{m1}+b_{m1})x_1 &+& (a_{m2}+b_{m2})x_2 &+\cdots+& (a_{mn}+b_{mn})x_n &=& y_m+z_m
\end{array}$$

Es natural decir entonces que la matriz del nuevo sistema lineal es

$$\begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n}\\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn}\end{bmatrix}$$

Esto motiva la siguiente definición:

Definición 7. (Suma de matrices)
Si A y B son matrices $m\times n$, su *suma* es la matriz $m\times n$ definida por la fórmula

$$[\,a_{ij}\,]_{m\times n}+[\,b_{ij}\,]_{m\times n}=[\,a_{ij}+b_{ij}\,]_{m\times n}$$

Observemos que la suma de matrices tiene sentido únicamente para matrices que tienen el mismo número de filas y de columnas.

Ejemplo 7.
Sumemos las siguientes matrices:

a) $$\begin{bmatrix} 5 & 1\\ 4 & 2\\ -5 & 12\end{bmatrix}+\begin{bmatrix} 3 & -1\\ 7 & 0\\ -5 & 1\end{bmatrix}=\begin{bmatrix} 5+3 & 1-1\\ 4+7 & 2+0\\ -5-5 & 12+1\end{bmatrix}=\begin{bmatrix} 8 & 0\\ 11 & 2\\ -10 & 13\end{bmatrix}$$

b) $$\begin{bmatrix} -3 & 0\\ 0 & -8\end{bmatrix}+\begin{bmatrix} 4 & 0\\ 0 & -1\end{bmatrix}=\begin{bmatrix} -3+4 & 0+0\\ 0+0 & -8-1\end{bmatrix}=\begin{bmatrix} 1 & 0\\ 0 & -9\end{bmatrix}$$

c) $$\begin{bmatrix} 2 & 5\\ 0 & -4\end{bmatrix}+\begin{bmatrix} -6 & 3\\ 0 & 8\end{bmatrix}=\begin{bmatrix} 2-6 & 5+3\\ 0+0 & -4+8\end{bmatrix}=\begin{bmatrix} -4 & 8\\ 0 & 4\end{bmatrix}$$

d) $$\begin{bmatrix} -1 & 2 & 6 & 4\\ 2 & 7 & 5 & 3\\ 4 & 3 & 0 & 8\end{bmatrix}+\begin{bmatrix} 1 & 2 & 0 & 3\\ 5 & 7 & -4 & 3\\ 1 & 0 & 4 & 6\end{bmatrix}=\begin{bmatrix} -1+1 & 2+2 & 6+0 & 4+3\\ 2+5 & 7+7 & 5-4 & 3+3\\ 4+1 & 3+0 & 0+4 & 8+6\end{bmatrix}$$

$$=\begin{bmatrix} 0 & 4 & 6 & 7\\ 7 & 14 & 1 & 6\\ 5 & 3 & 4 & 14\end{bmatrix}$$

b. Multiplicación de un escalar por una matriz

De manera similar a lo hecho para la suma, si inicialmente tenemos un sistema de ecuaciones lineales

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= y_2 \\ \vdots \qquad \vdots \qquad \vdots \qquad \vdots \quad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= y_m \end{aligned}$$

y multiplicamos sus coeficientes por un número k obtendremos el sistema

$$\begin{aligned} k\,a_{11}x_1 + k\,a_{12}x_2 + \cdots + k\,a_{1n}x_n &= k\,y_1 \\ k\,a_{21}x_1 + k\,a_{22}x_2 + \cdots + k\,a_{2n}x_n &= k\,y_2 \\ \vdots \qquad \vdots \qquad \vdots \qquad \vdots \quad &\quad \vdots \\ k\,a_{m1}x_1 + k\,a_{m2}x_2 + \cdots + k\,a_{mn}x_n &= k\,y_m \end{aligned}$$

cuya matriz sería entonces

$$\begin{bmatrix} k\,a_{11} & k\,a_{12} & \cdots & k\,a_{1n} \\ k\,a_{21} & k\,a_{22} & \cdots & k\,a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ k\,a_{m1} & k\,a_{m2} & \cdots & k\,a_{mn} \end{bmatrix}$$

Lo anterior justifica la siguiente definición:

Definición 8. (Multiplicación de un escalar por una matriz)
La *multiplicación de una matriz* $[\,a_{ij}\,]_{m\times n}$ *por un número (o escalar)* k está definida como

$$k\,[\,a_{ij}\,]_{m\times n} = [\,k\,a_{ij}\,]_{m\times n}$$

Ejemplo 8.

a) $$-3\begin{bmatrix} 5 & 1 \\ 4 & 2 \\ -5 & 12 \end{bmatrix} = \begin{bmatrix} -3(5) & -3(1) \\ -3(4) & -3(2) \\ -3(-5) & -3(12) \end{bmatrix} = \begin{bmatrix} -15 & -3 \\ -12 & -6 \\ 15 & -36 \end{bmatrix}$$

b) $$4\begin{bmatrix} 2 & 6 \\ 7 & 1 \end{bmatrix} - 5\begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 8 & 24 \\ 28 & 4 \end{bmatrix} + \begin{bmatrix} -10 & -5 \\ -15 & -10 \end{bmatrix} = \begin{bmatrix} -2 & 19 \\ 13 & -6 \end{bmatrix}$$

▲

Luego de haber definido la suma de matrices y su multiplicación por un escalar, surge la pregunta sobre qué propiedades algebraicas ya conocidas en los números se mantienen aún, dentro de este nuevo contexto. La respuesta la encontramos en el siguiente teorema:

Teorema 1. ***(Propiedades de la suma y multiplicación por un escalar)***
Sean A, B y C tres matrices $m \times n$. Entonces

a) $A + B = B + A$ *(ley conmutativa).*

b) $A + (B + C) = (A + B) + C$ *(ley asociativa).*

c) *Existe una matriz $m \times n$, que es la matriz cero (0), tal que $A + 0 = A$.*

d) *Para cada matriz $m \times n$, A, existe la matriz $m \times n$, $-A$, conformada por los negativos de las entradas de la matriz A, tal que $A + (-A) = 0$.*

Además, para todo par de escalares $k, l \in \mathbb{R}$ se tiene que:

e) $k(lA) = (kl)A$

f) $k(A + B) = kA + kB$

g) $(k + l)A = kA + lA$

Demostración

Todas las propiedades anteriores se reducen a propiedades bien conocidas de los números reales (Volumen 0 (Fundamentos)). Probemos el literal b) para ilustrar esto; los otros literales quedan como ejercicio (simple) para el lector.

Sean $A = [a_{ij}]_{m\times n}$, $B = [b_{ij}]_{m\times n}$, $C = [c_{ij}]_{m\times n}$; entonces, utilizando la propiedad asociativa de los números reales, se tiene que:

$$\begin{aligned}
A + (B + C) &= [a_{ij}] + ([b_{ij}] + [c_{ij}]) = [a_{ij}] + ([b_{ij} + c_{ij}]) \\
&= [a_{ij} + (b_{ij} + c_{ij})] \\
&= [(a_{ij} + b_{ij}) + c_{ij}] = [a_{ij} + b_{ij}] + [c_{ij}] \\
&= ([a_{ij}] + [b_{ij}]) + [c_{ij}] \\
&= (A + B) + C \quad \blacksquare
\end{aligned}$$

Ejemplo 9.

Recurriendo a las propiedades del ejercicio anterior, podemos escribir las siguientes operaciones matriciales en la forma indicada:

a) $7\begin{bmatrix} -3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix} - 2\begin{bmatrix} 4 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 6 \end{bmatrix} + \begin{bmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 3 \end{bmatrix} - \begin{bmatrix} 7 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -4 \end{bmatrix}$

$= \begin{bmatrix} -21 & 0 & 0 \\ 0 & 35 & 0 \\ 0 & 0 & 14 \end{bmatrix} + \begin{bmatrix} -8 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -12 \end{bmatrix} + \begin{bmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 3 \end{bmatrix} + \begin{bmatrix} -7 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 4 \end{bmatrix}$

$= \begin{bmatrix} -28 & 0 & 0 \\ 0 & 40 & 0 \\ 0 & 0 & 9 \end{bmatrix}$

b) $3\begin{bmatrix} -1 & 0 & 0 \\ 4 & -5 & 0 \\ 3 & 1 & -2 \end{bmatrix} + 6\begin{bmatrix} 6 & 0 & 0 \\ 3 & -9 & 0 \\ -7 & -4 & 1 \end{bmatrix} - 2\begin{bmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -5 \end{bmatrix} +$

$\begin{bmatrix} 3 & 8 & 5 \\ 2 & 6 & -10 \\ 15 & -4 & -11 \end{bmatrix}$

$= \begin{bmatrix} -3 & 0 & 0 \\ 12 & -15 & 0 \\ 9 & 3 & -6 \end{bmatrix} + \begin{bmatrix} 36 & 0 & 0 \\ 18 & -54 & 0 \\ -42 & -24 & 6 \end{bmatrix} + \begin{bmatrix} -12 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & 10 \end{bmatrix} +$

$\begin{bmatrix} 3 & 8 & 5 \\ 2 & 6 & -10 \\ 15 & -4 & -11 \end{bmatrix}$

$= \begin{bmatrix} 24 & 8 & 5 \\ 32 & -69 & -10 \\ -18 & -25 & -1 \end{bmatrix}$

c. Multiplicación de matrices

Ahora estudiemos la tercera operación sobre las matrices. Supongamos que tenemos un par de sistemas de ecuaciones lineales

$$\begin{aligned} a_{11}y_1 + a_{12}y_2 + \cdots + a_{1p}y_p &= z_1 \\ a_{21}y_1 + a_{22}y_2 + \cdots + a_{2p}y_p &= z_2 \\ \vdots \quad\quad \vdots \quad\quad \vdots \quad\quad \vdots \quad & \quad \vdots \\ a_{m1}y_1 + a_{m2}y_2 + \cdots + a_{mp}y_p &= z_m \end{aligned} \tag{2}$$

y

$$\begin{aligned} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n &= y_1 \\ b_{21}x_1 + b_{22}x_2 + \cdots + b_{2n}x_n &= y_2 \\ \vdots \qquad \vdots \qquad \vdots \qquad \vdots \quad &\ \ \vdots \\ b_{p1}x_1 + b_{p2}x_2 + \cdots + b_{pn}x_n &= y_p \end{aligned} \tag{3}$$

Entonces, para escribir $z_1, \ldots, z_m$ en términos de $x_1, \ldots, x_n$ debemos sustituir $y_1, \ldots, y_p$ en (2) por las respectivas expresiones en (3). Así obtenemos (después de algunos cálculos elementales que el lector podría corroborar) que

$$\begin{aligned} c_{11}x_1 + c_{12}x_2 + \cdots + c_{1n}x_n &= z_1 \\ c_{21}x_1 + c_{22}x_2 + \cdots + c_{2n}x_n &= z_2 \\ \vdots \qquad \vdots \qquad \vdots \qquad \vdots \quad &\ \ \vdots \\ c_{m1}x_1 + c_{m2}x_2 + \cdots + c_{mn}x_n &= z_m \end{aligned} \tag{4}$$

donde

$$c_{ij} = a_{i1}b_{1j} + \cdots + a_{ip}b_{pj} \qquad \text{para} \quad i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, p$$

y este será el origen de la siguiente definición:

Definición 9. (Multiplicación de dos matrices)
La *multiplicación* o el *producto* de una matriz $m \times p$, A, por una matriz $p \times n$, B, es una matriz $m \times n$, C, cuya entrada c_{ij} está definida por

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}$$

para cada $1 \leq i \leq m$ y $1 \leq j \leq n$. En este caso, escribiremos $C = AB$.

A continuación se muestran los elementos que son considerados en la matriz A y en la matriz B con el fin de obtener la entrada c_{ij}:

$$\begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \vdots & \vdots \\ \vdots & \boldsymbol{c_{ij}} & \vdots \\ \vdots & \vdots & \vdots \\ c_{m1} & \cdots & c_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & \cdots & \cdots & a_{1p} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \boldsymbol{a_{i1}} & \cdots & \boldsymbol{a_{ik}} & \cdots & \boldsymbol{a_{ip}} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{m1} & \cdots & \cdots & \cdots & a_{mp} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & \boldsymbol{b_{1j}} & \cdots & b_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \vdots & \cdots & \boldsymbol{b_{kj}} & \cdots & \vdots \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ b_{p1} & \cdots & \boldsymbol{b_{pj}} & \cdots & b_{pn} \end{bmatrix}$$

Observemos que en la multiplicación de matrices el número de columnas de la primera matriz *debe coincidir* con el número de filas de la segunda matriz.

Ejemplo 10.
Hallemos AB si

a) $A = \begin{bmatrix} 2 & -1 \\ 3 & 4 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & 1 & 5 \\ -2 & 1 & 1 \end{bmatrix}$

b) $A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$ y $B = \begin{bmatrix} 5 & 7 \\ -2 & 3 \end{bmatrix}$

Solución

a) Puesto que A es una matriz de tamaño 2×2 y B es una matriz de tamaño 2×3, el producto AB está bien definido y es una matriz de tamaño 2×3. Para obtener la primera fila de la matriz producto AB, multiplicamos los términos respectivos de la primera fila, $[2 \ \ -1]$, de A por cada uno de los respectivos términos de cada una de las columnas de B

$$\begin{bmatrix} 0 \\ -2 \end{bmatrix} \quad , \quad \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 5 \\ 1 \end{bmatrix}$$

respectivamente. Esto es,

$$AB = \begin{bmatrix} (2)(0) + (-1)(-2) & (2)(1) + (-1)(1) & (2)(5) + (-1)(1) \\ \bullet & \bullet & \bullet \end{bmatrix}$$

Para obtener la segunda fila de AB, multiplicamos los términos de la segunda fila, $[3 \quad 4]$, de A por los términos de cada una de las columnas de B. Así,

$$AB = \begin{bmatrix} 2 & 1 & 9 \\ (3)(0) + (4)(-2) & (3)(1) + (4)(1) & (3)(5) + (4)(1) \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 1 & 9 \\ -8 & 7 & 19 \end{bmatrix}$$

Observemos que para estas matrices no está definido el producto BA debido a la incompatibilidad de sus tamaños, es decir, que el número de columnas de B es distinto al número de filas de A.

b) Como A y B son matrices 2×2, el producto AB está bien definido y es una matriz 2×2. La primera fila de AB se obtiene multiplicando los términos de la primera fila de A por los términos de las columnas de B. Así,

$$AB = \begin{bmatrix} 5+2 & 7-3 \\ \bullet & \bullet \end{bmatrix}$$

La segunda fila de AB se obtiene multiplicando los términos de la segunda fila de A por los términos de cada una de las columnas de B. Por tanto,

$$AB = \begin{bmatrix} 7 & 4 \\ 0-4 & 0+6 \end{bmatrix} \quad = \quad \begin{bmatrix} 7 & 4 \\ -4 & 6 \end{bmatrix}$$

▲

Ahora: contrario a lo que tal vez esperaría un lector desprevenido, el siguiente ejemplo muestra que la multiplicación de matrices es *no-conmutativa*; es decir, en general, los productos AB y BA de matrices no son necesariamente iguales. Pero esto no es sorpresivo, pues, al fin y al cabo, el proceso llevado a cabo en (2), (3) y (4) arriba no es más que una forma de composición de funciones, y ya sabemos (Volumen 0 (Fundamentos)) que la composición de funciones no es, en general, conmutativa.

Ejemplo 11. (En general, AB $\neq$ BA)
Calculemos AB y BA si

$$A = \begin{bmatrix} 1 & -1 \\ 3 & -3 \\ 2 & 8 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & -1 & -5 \\ -2 & 3 & -1 \end{bmatrix}$$

Solución

Como A es una matriz 3×2 y B es una matriz 2×3, entonces los productos AB y BA están bien definidos. Para obtener la primera fila de la matriz AB, multiplicamos la primera fila $[1 \;\; -1]$ de A por cada una de las columnas de B:

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix} \;,\; \begin{bmatrix} -1 \\ 3 \end{bmatrix} \;\text{y}\; \begin{bmatrix} -5 \\ -1 \end{bmatrix}$$

respectivamente. Esto es,

$$AB = \begin{bmatrix} 1+2 & -1-3 & -5+1 \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{bmatrix}$$

Para obtener la segunda fila de AB, multiplicamos la segunda fila $[3 \quad -3]$ de A por cada una de las columnas de B. Así,

$$AB = \begin{bmatrix} 3 & -4 & -4 \\ 3+6 & -3-9 & -15+3 \\ \bullet & \bullet & \bullet \end{bmatrix}$$

Para obtener la tercera fila de AB, multiplicamos la tercera fila $[2 \quad 8]$ de A por cada una de las columnas de B. Así,

$$AB = \begin{bmatrix} 3 & -4 & -4 \\ 9 & -12 & -12 \\ 2-16 & -2+24 & -10-8 \end{bmatrix}$$

Por tanto,

$$AB = \begin{bmatrix} 3 & -4 & -4 \\ 9 & -12 & -12 \\ -14 & 22 & -18 \end{bmatrix}$$

De forma similar, para obtener la primera fila de BA, multiplicamos la primera fila $[1 \quad -1 \quad -\ 5]$ de B por cada una de las columnas de A. Así,

$$BA = \begin{bmatrix} 1-3-10 & -1+3-40 \\ \bullet & \bullet \end{bmatrix}$$

y para obtener la segunda fila de BA, multiplicamos la segunda fila $[-2 \quad 3 \quad -1]$ de B por cada una de las columnas de A. Así,

$$BA = \begin{bmatrix} -12 & -38 \\ -2+9-2 & 2-9-8 \end{bmatrix}$$

Por tanto,

$$BA = \begin{bmatrix} -12 & -38 \\ 5 & -15 \end{bmatrix}$$

Observemos entonces que, no sólo $AB \neq BA$, sino que las matrices AB y BA son de diferente tamaño. ▲

Y arribamos, entonces, a describir cuáles de las propiedades de los números todavía preservan las matrices, y su particular definición de multiplicación.

Teorema 2. ***(Propiedades de la multiplicación de matrices)***
Supongamos que A, B y C son matrices compatibles (es decir, que las multiplicaciones pueden realizarse) para la multiplicación. Entonces

a) $AB \neq BA$ *(no se satisface la ley conmutativa)*

b) $A(BC) = (AB)C$ *(ley asociativa)*

c) $A(B + C) = AB + AC$ *(ley distributiva a derecha)*

d) $(B + C)A = BA + CA$ *(ley distributiva a izquierda)*

e) $k(AB) = (kA)B = A(kB)$, *donde* $k \in \mathbb{R}$

f) *Para toda matriz cuadrada* A *de tamaño* $n \times n$ *se tiene que* $AI_n = I_nA = A$.

Demostración

a) En el ejemplo 11 anterior mostramos que si $A = \begin{bmatrix} 1 & -1 \\ 3 & -3 \\ 2 & 8 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & -1 & -5 \\ -2 & 3 & -1 \end{bmatrix}$ entonces $AB \neq BA$.

b) Si $A = [a_{ij}]$, $B = [b_{ij}]$, $C = [c_{ij}]$, entonces

i) $BC = [d_{ij}]$, donde $d_{ij} = \sum\limits_{k=1}^{n} b_{ik}\, c_{kj}$.

ii) $AB = [e_{ij}]$, donde $e_{ij} = \sum\limits_{k=1}^{n} a_{ik}\, b_{kj}$.

iii) $A(BC) = [f_{ij}]$, donde $f_{ij} = \sum\limits_{k=1}^{n} a_{ik}\, d_{kj} = \sum\limits_{k=1}^{n} a_{ik} \sum\limits_{l=1}^{n} b_{kl}\, c_{lj}$.

iv) $(AB)C = [g_{ij}]$, donde $g_{ij} = \sum\limits_{k=1}^{n} e_{ik}\, c_{kj} = \sum\limits_{k=1}^{n} [\sum\limits_{l=1}^{n} a_{il}\, b_{lk}]\, c_{kj} = \sum\limits_{l=1}^{n} a_{il} \sum\limits_{k=1}^{n} b_{lk} c_{kj}$.

Un cambio de índices nos muestra inmediatamente que $g_{ij} = f_{ij}$ y, por tanto, $A(BC) = (AB)C$.

Las demostraciones de los literales restantes son similares y se dejan como ejercicios para el lector. ■

Ejemplo 12. (A(B+C)=AB+AC)
Consideremos las siguientes matrices:

$$A = \begin{bmatrix} 1 & 2 & -3 \\ -2 & 4 & 6 \\ 1 & 2 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Entonces

$$B + C = \begin{bmatrix} -1 & 0 & 0 \\ 4 & 2 & -6 \\ 0 & -2 & 1 \end{bmatrix}$$

y

$$A(B + C) = \begin{bmatrix} 7 & 10 & -15 \\ 18 & -4 & -18 \\ 7 & -6 & -7 \end{bmatrix}$$

Además,

$$AB = \begin{bmatrix} 5 & 10 & -15 \\ 6 & -12 & -18 \\ -3 & -6 & -15 \end{bmatrix}$$

y

$$AC = \begin{bmatrix} 2 & 0 & 0 \\ 12 & 8 & 0 \\ 10 & 0 & 8 \end{bmatrix}$$

lo que implica que

$$AB + AC = \begin{bmatrix} 7 & 10 & -15 \\ 18 & -4 & -18 \\ 7 & -6 & -7 \end{bmatrix}$$

Por tanto, $A(B + C) = AB + AC$.

Ejemplo 13. (A(BC)=(AB)C)
Con las matrices A, B, C del ejemplo anterior, mostremos que, también, $A(BC) = (AB)C$.

Solución

Aquí,

$$BC = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2 & 3 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & 6 & -9 \\ -2 & -3 & 0 \\ -5 & 0 & -3 \end{bmatrix}$$

Y así,

$$A(BC) = \begin{bmatrix} 1 & 2 & -3 \\ -2 & 4 & 6 \\ 1 & 2 & 5 \end{bmatrix} \begin{bmatrix} -1 & 6 & -9 \\ -2 & -3 & 0 \\ -5 & 0 & -3 \end{bmatrix}$$

$$= \begin{bmatrix} 10 & 0 & 0 \\ -36 & -24 & 0 \\ -30 & 0 & -24 \end{bmatrix}$$

Ahora: tomando la matriz AB del ejercicio anterior se tiene que:

$$(AB)C = \begin{bmatrix} 5 & 10 & -15 \\ 6 & -12 & -18 \\ -3 & -6 & -15 \end{bmatrix} \begin{bmatrix} 1 & -2 & 3 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 0 & 0 \\ -36 & -24 & 0 \\ -30 & 0 & -24 \end{bmatrix}$$

Ejemplo 14. ($\mathbf{AI_n = I_nA = A}$)
Calculemos AI_3 e I_3A si I_3 es la matriz identidad 3×3 y

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 10 & 12 & 20 \\ 11 & 0 & 5 \end{bmatrix} \quad \text{y} \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Solución

$$AI_3 = \begin{bmatrix} 2 & -3 & 1 \\ 10 & 12 & 20 \\ 11 & 0 & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -3 & 1 \\ 10 & 12 & 20 \\ 11 & 0 & 5 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -3 & 1 \\ 10 & 12 & 20 \\ 11 & 0 & 5 \end{bmatrix} = I_3A$$

Ejercicios 3

1) Considere las siguientes matrices:

$$A = \begin{bmatrix} -4 & 5 & 1 \\ -2 & 3 & \frac{1}{2} \\ \frac{1}{3} & 0 & 8 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & -6 & 5 \\ 7 & 2 & -1 \\ 1 & 3 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 2 & 0 \\ 4 & 1 & -3 \end{bmatrix}$$

Calcule, si es posible, las siguientes operaciones entre matrices:

a) $A + B$	b) $8A$
c) $B - A$	d) $3B - 2A$
e) $C - A$	f) $2C - A - 2B$
g) $B + 3C$	h) $\frac{1}{2}A + \frac{3}{2}C - \frac{1}{3}B$

donde $B - A \equiv B + (-A)$; $C - A \equiv C + (-A)$; y $-A \equiv (-1)A$; etc.

2) Considere las siguientes matrices:

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 5 & -1 \\ 3 & 3 \end{bmatrix}$$

$$C = \begin{bmatrix} 8 & 7 & 3 \\ 1 & -4 & 6 \end{bmatrix} \qquad D = \begin{bmatrix} 3 & 5 & -6 & 4 \\ 1 & 2 & 0 & 1 \end{bmatrix}$$

Calcule, si es posible, las siguientes operaciones entre matrices:

a) AB b) BA c) AD
d) BC e) DA f) ABC

3) Considere las siguientes matrices:

a) $A = \begin{bmatrix} -10 & -8 & 4 & 1 \\ -11 & 6 & 0 & 3 \\ 0 & 1 & 2 & 7 \end{bmatrix}$ b) $B = \begin{bmatrix} 7 & 9 & 9 \\ 14 & 6 & 1 \end{bmatrix}$

c) $C = \begin{bmatrix} 3 & 0 & 8 \\ 0 & 4 & 0 \\ 8 & 0 & 10 \end{bmatrix}$ d) $D = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ 0 & 0 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{5} & \frac{1}{6} & 0 & \frac{1}{3} \\ \frac{3}{7} & \frac{3}{4} & \frac{5}{6} & 1 \end{bmatrix}$

e) $E = \begin{bmatrix} 12 & -6 & -7 & -1 \\ 4 & 5 & 5 & 3 \\ 2 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 \end{bmatrix}$ f) $F = \begin{bmatrix} 1 & -5 & -6 \\ 4 & 0 & -3 \end{bmatrix}$

Calcule, si es posible, las siguientes operaciones matriciales:

a) $BC - 5F$ b) BAD
c) $4AD + 3E$ d) $4BC$

4) Pruebe que si

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 6 & -11 & 6 \end{bmatrix}$$

entonces $A^3 - 6A^2 + 11A - 6I = 0$ donde $A^3 = A \cdot A \cdot A$; $A^2 = A \cdot A$ y $I = I_3$.

4. Otros tipos de matrices

Además de las matrices básicas que hemos definido (nula, idéntica, etc.), existen otros tipos de matrices que surgen muy naturalmente cuando estudiamos problemas lineales. De hecho, ya aparecieron cuando estudiábamos el método gaussiano de solución a sistemas lineales. Allí, además, describíamos la noción de *operación fila*, que ahora recordamos para efectos de exposición.

Definición 10. (Operaciones fila)
Sea A una matriz $m \times n$. Las *operaciones fila* sobre A son:

a) Multiplicar la i-ésima fila de A por un escalar diferente de cero.

b) Sumar un múltiplo de la i-ésima fila de A a la j-ésima fila de A.

c) Intercambiar la i-ésima y la j-ésima filas de A.

Definición 11. (Matriz elemental)
Una matriz cuadrada de tamaño n es una *matriz elemental* si es obtenida a partir de la matriz idéntica de tamaño n mediante una sola operación fila.

Ejemplo 15.
Veamos que las siguientes seis matrices son elementales:

a) $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}$ b) $B = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

c) $C = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ d) $D = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix} \quad k \neq 0$

e) $E = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \quad k \neq 0$ f) $F = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Solución

a) A es una matriz elemental porque se obtiene de I_4 multiplicando por 6 su cuarta fila.

b) B es una matriz elemental porque se obtiene de I_3 multiplicando por 4 la primera fila y sumándola a la segunda.

c) C es una matriz elemental porque se obtiene de I_3 intercambiando la primera y la segunda filas.

d) D es una matriz elemental porque se obtiene de I_2 multiplicando su primera fila por k.

e) E es una matriz elemental porque se obtiene de I_2 multiplicando por k la segunda fila y sumándola a la primera.

f) F es una matriz elemental porque se obtiene de I_2 intercambiando la primera y la segunda filas. ▲

Teorema 3. ***(Operaciones fila y multiplicación de matrices)***
Cada operación fila sobre una matriz A puede realizarse multiplicando a la izquierda de la matriz A por la correspondiente matriz elemental.

Demostración

Sea

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

a) Si multiplicamos la fila i de la matriz A por $k \neq 0$, obtendremos:

$$A^* = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ ka_{i1} & ka_{i2} & \cdots & ka_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Y observemos que:

$$\text{Fila i} \longrightarrow \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & \cdots & k & \cdots & 0 \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ ka_{i1} & ka_{i2} & \cdots & ka_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = A^*$$

b) Si sumamos la fila j de la matriz A a la fila i, obtendremos:

$$A^* = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1}+a_{j1} & a_{i2}+a_{j2} & \cdots & a_{in}+a_{jn} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{matrix} \\ \\ \longleftarrow \text{Fila j} \\ \\ \longleftarrow \text{Fila i} \\ \\ \\ \end{matrix}$$

Y observemos que:

$$\begin{matrix} \\ \\ \text{Fila j} \longrightarrow \\ \\ \text{Fila i} \longrightarrow \\ \\ \\ \end{matrix} \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1}+a_{j1} & a_{i2}+a_{j2} & \cdots & a_{in}+a_{jn} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = A^*$$

c) Si intercambiamos la fila i con la fila j, obtendremos:

$$A^* = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{array}{l} \\ \\ \longleftarrow \text{Fila j} \\ \\ \longleftarrow \text{Fila i} \\ \\ \\ \end{array}$$

y notemos que:

$$\begin{array}{r} \\ \\ \\ \text{Fila j} \longrightarrow \\ \\ \text{Fila i} \longrightarrow \\ \\ \\ \end{array} \begin{bmatrix} 1 & 0 & \cdots & \mathbf{0} & 0 & \cdots & \mathbf{0} & \cdots & 0 \\ 0 & 1 & \cdots & \mathbf{0} & 0 & \cdots & \mathbf{0} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \mathbf{0} & 0 & \cdots & \mathbf{1} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \mathbf{1} & 0 & \cdots & \mathbf{0} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \mathbf{0} & 0 & \cdots & \mathbf{0} & \cdots & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} =$$

$$\begin{array}{r} \\ \\ \\ \text{Fila j} \longrightarrow \\ \\ \text{Fila i} \longrightarrow \\ \\ \\ \end{array} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = A^* \quad \blacksquare$$

Ejemplo 16.

a) Sean

$$A = \begin{bmatrix} 2 & 6 & 5 \\ 4 & 1 & 3 \\ 1 & 1 & 2 \\ 7 & 3 & 4 \end{bmatrix}$$

y B la matriz que se obtiene de A multiplicando por 3 la tercera fila; es decir,

$$B = \begin{bmatrix} 2 & 6 & 5 \\ 4 & 1 & 3 \\ 3 & 3 & 6 \\ 7 & 3 & 4 \end{bmatrix}$$

Observemos que B puede obtenerse multiplicando a izquierda de la matriz A por la matriz elemental que se obtiene de I_4 multiplicando por 3 la tercera fila:

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 6 & 5 \\ 4 & 1 & 3 \\ 1 & 1 & 2 \\ 7 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 6 & 5 \\ 4 & 1 & 3 \\ 3 & 3 & 6 \\ 7 & 3 & 4 \end{bmatrix}$$

b) Sean

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 10 & 12 & 20 \\ 11 & 0 & 5 \end{bmatrix}$$

y B la matriz que se obtiene de A multiplicando por 4 la primera fila y sumándola a la segunda; es decir,

$$B = \begin{bmatrix} 2 & -3 & 1 \\ 18 & 0 & 24 \\ 11 & 0 & 5 \end{bmatrix}$$

Observemos que B puede obtenerse multiplicando a izquierda de la matriz A por la matriz elemental que se obtiene de I_3 multiplicando por 4 la primera fila y sumándola a la segunda:

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -3 & 1 \\ 10 & 12 & 20 \\ 11 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 2 & -3 & 1 \\ 18 & 0 & 24 \\ 11 & 0 & 5 \end{bmatrix}$$

c) Sean

$$A = \begin{bmatrix} 5 & 1 \\ 4 & 2 \\ 5 & 9 \end{bmatrix}$$

y B la matriz que se obtiene de A intercambiando la primera y segunda filas; es decir,

$$B = \begin{bmatrix} 4 & 2 \\ 5 & 1 \\ 5 & 9 \end{bmatrix}$$

Observemos que B puede obtenerse multiplicando a izquierda de la matriz A por la matriz elemental que se obtiene de I_3 intercambiando la primera y segunda filas:

$$B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & 1 \\ 4 & 2 \\ 5 & 9 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 5 & 1 \\ 5 & 9 \end{bmatrix} \blacktriangle$$

Otro de los conceptos que será siempre útil y recurrente en el estudio de problemas que involucren matrices, es el siguiente:

Definición 12. (Traspuesta de una matriz)
La *traspuesta de una matriz* $A = [\, a_{ij} \,]_{m\times n}$, que denotaremos por A^T, se obtiene de la matriz A escribiendo cada una de sus filas como columnas preservando el orden. Es decir, la primera fila de A corresponderá a la primera columna de A^T, la segunda fila de A corresponderá a la segunda columna de A^T, y así sucesivamente. Luego la matriz A^T es de orden $n \times m$. Así, $A^T = [\, a_{ji} \,]_{n\times m}$.

Ejemplo 17.
Hallemos la traspuesta de las siguientes matrices:

a) $A = \begin{bmatrix} 1 & -1 \\ 3 & -3 \\ 2 & 8 \end{bmatrix}$

b) $A = \begin{bmatrix} 0 & -3 & 2 \end{bmatrix}$

c) $A = \begin{bmatrix} 1 & 4 & 5 & -6 \\ 3 & 2 & 1 & 0 \\ -1 & -3 & -2 & -8 \end{bmatrix}$

d) $A = \begin{bmatrix} 7 & 9 & 10 \\ 25 & -1 & 18 \end{bmatrix}$

Solución

a) $A^T = \begin{bmatrix} 1 & -1 \\ 3 & -3 \\ 2 & 8 \end{bmatrix}^T = \begin{bmatrix} 1 & 3 & 2 \\ -1 & -3 & 8 \end{bmatrix}$

b) $A^T = \begin{bmatrix} 0 & -3 & 2 \end{bmatrix}^T = \begin{bmatrix} 0 \\ -3 \\ 2 \end{bmatrix}$

c) $A^T = \begin{bmatrix} 1 & 4 & 5 & -6 \\ 3 & 2 & 1 & 0 \\ -1 & -3 & -2 & -8 \end{bmatrix}^T = \begin{bmatrix} 1 & 3 & -1 \\ 4 & 2 & -3 \\ 5 & 1 & -2 \\ -6 & 0 & -8 \end{bmatrix}$

d) $A^T = \begin{bmatrix} 7 & 9 & 10 \\ 25 & -1 & 18 \end{bmatrix}^T = \begin{bmatrix} 7 & 25 \\ 9 & -1 \\ 10 & 18 \end{bmatrix}$

Teorema 4. ***(Propiedades de la traspuesta de una matriz)***

a) $I_n^T = I_n$

b) $(A + B)^T = A^T + B^T$

c) $(kA)^T = kA^T, \; k \in \mathbb{R}$

d) $(AB)^T = B^T A^T$

Demostración

a) Sea $I_n = [\delta_{ij}]_{n\times n}$, donde $\delta_{ij} = 1$ si $i = j$, y $\delta_{ij} = 0$ si $i \neq j$; entonces

$$I_n^T = [\delta_{ij}]_{n\times n}^T = [\delta_{ji}]_{n\times n} = I_n$$

b) Sean $A = [a_{ij}]_{m\times n}$ y $B = [b_{ij}]_{m\times n}$; entonces

$$\begin{aligned}(A + B)^T &= [a_{ij} + b_{ij}]_{m\times n}^T = [a_{ji} + b_{ji}]_{n\times m} \\ &= [a_{ji}]_{n\times m} + [b_{ji}]_{n\times m} = A^T + B^T\end{aligned}$$

c) Si $A = [a_{ij}]_{m\times n}$ y $k \in \mathbb{R}$, entonces

$$(kA)^T = [ka_{ji}]_{n\times m} = k[a_{ji}]_{n\times m} = kA^T$$

d) Si $A = [a_{ij}]_{m\times p}$ y $B = [b_{ij}]_{p\times n}$; entonces

i) $(AB)^T = [c_{ji}]_{n\times m}$, donde $c_{ij} = \sum\limits_{k=1}^{p} a_{ik}\, b_{kj}$.

ii) $B^T A^T = [b_{ji}]_{n\times p}[a_{ji}]_{p\times m} = [d_{ji}]$, donde $d_{ji} = \sum\limits_{k=1}^{p} a_{jk}\, b_{ki}$. Observemos que $c_{ji} = \sum\limits_{k=1}^{p} a_{jk}\, b_{ki} = d_{ji}$. Luego, $(AB)^T = B^T A^T$ ■

Ejemplo 18. ($\mathbf{(AB)^T = B^T A^T}$)
Con las matrices

$$A = \begin{bmatrix} 1 & 2 & -3 \\ -2 & 4 & 6 \\ 1 & 2 & 5 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$$

Comprobemos que $(AB)^T = B^T A^T$

Solución

Aquí,

$$AB = \begin{bmatrix} 5 & 10 & -15 \\ 6 & -12 & -18 \\ -3 & -6 & -15 \end{bmatrix}$$

y así,

$$(AB)^T = \begin{bmatrix} 5 & 6 & -3 \\ 10 & -12 & -6 \\ -15 & -18 & -15 \end{bmatrix}$$

Ahora:

$$B^T = \begin{bmatrix} -2 & 2 & -1 \\ 2 & 1 & -2 \\ -3 & -6 & 0 \end{bmatrix} \quad \text{y} \quad A^T = \begin{bmatrix} 1 & -2 & 1 \\ 2 & 4 & 2 \\ -3 & 6 & 5 \end{bmatrix}$$

y así,

$$B^T A^T = \begin{bmatrix} 5 & 6 & -3 \\ 10 & -12 & -6 \\ -15 & -18 & -15 \end{bmatrix}$$

Ejemplo 19. (Un proceso elemental de Markov)
El estado del uso del suelo en una ciudad de 50 millas cuadradas de superficie en el año 1993 fue

$$\text{I (Uso residencial) } 30\,\%$$

$$\text{II (Uso comercial) } 20\,\%$$

$$\text{III (Uso industrial) } 50\,\%$$

Encontremos los estados en 1998 y 2003, suponiendo que las "probabilidades de transición" para intervalos de cinco años están dadas por la siguiente matriz:

$$A \equiv \begin{array}{c|ccc} & \text{a: I} & \text{a: II} & \text{a: III} \\ \text{De: I} & 0.8 & 0.1 & 0.1 \\ \text{De: II} & 0.1 & 0.7 & 0.2 \\ \text{De: III} & 0 & 0.1 & 0.9 \end{array}$$

Aquí, por ejemplo, 0.8 es la probabilidad de que 1 milla que era de uso residencial se mantenga así al cabo de cinco años. En cambio, 0.2 es la probabilidad de que 1 milla que era de uso comercial, pase, en cinco años, a uso industrial.

Solución

A partir de A y del estado del año 1993 podemos calcular el estado en el año 1998: Sea $x = (\,30\,\%, 20\,\%, 50\,\%\,)$ el estado de 1993 y $y = (\,y_1, y_2, y_3\,)$ el estado de 1998; entonces $y^T = Ax^T$; es decir,

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0.8 & 0.1 & 0 \\ 0.1 & 0.7 & 0.1 \\ 0.1 & 0.2 & 0.9 \end{bmatrix} \begin{bmatrix} 30\,\% \\ 20\,\% \\ 50\,\% \end{bmatrix} = \begin{bmatrix} 26\,\% \\ 22\,\% \\ 52\,\% \end{bmatrix}$$

De manera similar, si $z = (\,z_1, z_2, z_3\,)$ es el estado en el año 2003, entonces $z^T = Ay^T = A(\,Ax^T\,) = A^2x^T$. Luego $z = (\,23\,\%, 23.2\,\%, 53.8\,\%\,)$. ¿Podría el lector estimar, de la misma forma, la distribución del suelo en el año 2012 y en el 2020? ▲

Quizás uno de los más importantes tipos de matrices que aparecen muy comúnmente en estudios de mecánica clásica (física) es el de *matriz simétrica*. Veamos su definición y propiedades.

Definición 13. (Matriz simétrica)
Una matriz cuadrada A de tamaño n es *simétrica* si $A = A^T$.

Ejemplo 20.
Las siguientes matrices son simétricas, pues coinciden con sus respectivas traspuestas:

a) $\begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & -8 \\ 1 & -8 & 5 \end{bmatrix}$ b) $\begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 4 & -5 & 10 \\ 4 & -1 & -6 & 2 \\ -5 & -6 & 1 & 3 \\ 10 & 2 & 3 & -1 \end{bmatrix}$ d) $\begin{bmatrix} 1 & 4 \\ 4 & 1 \end{bmatrix}$

Teorema 5. ***(Propiedades de las matrices simétricas)***
Sean A y B matrices cuadradas. Entonces:

a) Si A y B son simétricas, también $A + B$ es simétrica.

b) Si A es simétrica y $k \in \mathbb{R}$, también kA es simétrica.

c) Si A y B son simétricas y conmutan (es decir, $AB = BA$), entonces AB es simétrica.

Demostración

a) $(A+B)^T = A^T + B^T = A + B$

b) $(kA)^T = kA^T = kA$

c) $(AB)^T = B^T A^T = BA = AB$ ■

Ejemplo 21. (El producto de matrices simétricas no es siempre simétrico)
Consideremos las matrices simétricas

$$A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & -1 & 4 \\ -3 & 4 & 0 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} -1 & 2 & 1 \\ 2 & 3 & -1 \\ 1 & -1 & 2 \end{bmatrix}$$

y mostremos que AB *no* es simétrica.

Solución

En efecto,

$$\begin{aligned} AB &= \begin{bmatrix} 1 & 2 & -3 \\ 2 & -1 & 4 \\ -3 & 4 & 0 \end{bmatrix} \begin{bmatrix} -1 & 2 & 1 \\ 2 & 3 & -1 \\ 1 & -1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 11 & -7 \\ 0 & -3 & 11 \\ 11 & 6 & -7 \end{bmatrix} \end{aligned}$$

no es simétrica. ¿Será que BA es simétrica? ▲

Un concepto que también es útil es el de *traza* de una matriz:

Definición 14. (Traza de una matriz)
Si $A = [\, a_{ij} \,]_{n\times n}$ es una matriz cuadrada, entonces definimos la *traza* de A por

$$Tr(A) = \sum_{i=1}^{n} a_{ii}$$

En otras palabras, la traza es la suma de los elementos de la diagonal principal de la matriz.

Ejemplo 22.
Calculemos la traza de las siguientes matrices cuadradas:

a) $A = \begin{bmatrix} 4 & 1 \\ -2 & -5 \end{bmatrix}$ b) $B = \begin{bmatrix} 7 & 6 & 4 \\ 3 & -2 & -1 \\ 3 & 5 & 6 \end{bmatrix}$

Solución

a) $Tr(A) = 4 + (-5) = -1$ b) $Tr(B) = 7 + (-2) + 6 = 11$

Teorema 6. ***(Propiedades de la traza)***

a) $Tr(I_n) = n;\ Tr(0) = 0$

Si A *y* B *son matrices cuadradas de tamaño* n*, y* k *es un escalar, entonces*

b) $Tr(A + B) = Tr(A) + Tr(B)$

c) $Tr(kA) = kTr(A)$

d) $Tr(AB) = Tr(BA)$

e) En general, $Tr(AB) \neq Tr(A)Tr(B)$

Demostración

a) Se deja como ejercicio elemental para el lector.

b) $Tr(A + B) = \sum\limits_{i=1}^{n} (a_{ii} + b_{ii}) = \sum\limits_{i=1}^{n} a_{ii} + \sum\limits_{i=1}^{n} b_{ii} = Tr(A) + Tr(B)$

c) $Tr(kA) = \sum\limits_{i=1}^{n} ka_{ii} = k \sum\limits_{i=1}^{n} a_{ii} = kTr(A)$

d) $Tr(AB) = \sum\limits_{i=1}^{n} \left(\sum\limits_{k=1}^{n} (a_{ik})(b_{ki}) \right) = \sum\limits_{k=1}^{n} \left(\sum\limits_{i=1}^{n} (b_{ik})(a_{ki}) \right) = Tr(BA)$

e) El ejemplo 21 nos muestra dos matrices A y B tales que

$$Tr(AB) \neq Tr(A)Tr(B)$$

pues $Tr(AB) = -10$ pero $Tr(A) = 0$, $Tr(B) = 4$. ■

Ejemplo 23.
Consideremos las siguientes matrices cuadradas:

$$A = \begin{bmatrix} 1 & 3 & 3 \\ 2 & 4 & 1 \\ 6 & 5 & 2 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 2 & 4 & 8 \\ 1 & 5 & 0 \\ 6 & 2 & 1 \end{bmatrix}$$

Verifiquemos las propiedades de la traza.

Solución

Observemos que

$$A+B=\begin{bmatrix}3 & 7 & 11\\ 3 & 9 & 1\\ 12 & 7 & 3\end{bmatrix} \qquad 3A=\begin{bmatrix}3 & 9 & 9\\ 6 & 12 & 3\\ 18 & 15 & 6\end{bmatrix}$$

$$AB=\begin{bmatrix}23 & 25 & 11\\ 14 & 30 & 17\\ 29 & 53 & 50\end{bmatrix} \qquad BA=\begin{bmatrix}58 & 62 & 26\\ 11 & 23 & 8\\ 16 & 31 & 22\end{bmatrix}$$

Por tanto,

a) $Tr(A)+Tr(B)=7+8=15=Tr(A+B)$

b) $Tr(3A)=21=3Tr(A)$

c) $Tr(AB)=103=Tr(BA)$

d) $Tr(A)Tr(B)=56\neq 103=Tr(AB)$.

a. Matrices particionadas

Puesto que una matriz es una disposición rectangular de elementos, es posible dividirla en ordenaciones más pequeñas denominadas *submatrices*. Por ejemplo, podríamos escribir la matriz

$$A=\begin{bmatrix}a_{11} & a_{12} & a_{13} & a_{14}\\ a_{21} & a_{22} & a_{23} & a_{24}\\ a_{31} & a_{32} & a_{33} & a_{34}\end{bmatrix} \quad \text{como} \quad A=\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}$$

donde

$$A_{11}=\begin{bmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\end{bmatrix} \quad ; \qquad A_{12}=\begin{bmatrix}a_{14}\\ a_{24}\end{bmatrix}$$

$$A_{21}=\begin{bmatrix}a_{31} & a_{32} & a_{33}\end{bmatrix} \quad ; \qquad A_{22}=\begin{bmatrix}a_{34}\end{bmatrix}$$

Decimos, en este caso, que la matriz A es una *matriz particionada.*

Las operaciones fundamentales de suma y multiplicación continúan siendo, afortunadamente, aplicables a las matrices particionadas, aunque las matrices han de partirse de modo adecuado para poder realizar estas operaciones. Por ejemplo, si B es también de tamaño 3×4 y la partimos de manera similar a como lo hicimos con la matriz A anterior, obtenemos que

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

donde B_{ij} es del mismo tamaño que A_{ij}, y entonces podemos definir la suma $A + B$ como

$$A + B = \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{bmatrix}$$

Para el caso de la multiplicación, la única condición requerida es que el número de columnas de A_{ij} sea igual al número de filas de B_{jk} para todos los pares (i, j). Por ejemplo, el producto AB para las matrices A y B definidas arriba es

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}$$

El análisis de los elementos en los productos de las submatrices indica que el mismo resultado se habría obtenido mediante la multiplicación directa de las matrices originales sin partir. Por tanto, las submatrices pueden ser tratadas aquí como elementos ordinarios siempre que exista concordancia en la partición.

Ejemplo 24.
Consideremos las siguientes matrices A y B:

$$A = \begin{bmatrix} 3 & -2 & 3 & -1 \\ 0 & 5 & -1 & -3 \\ -2 & 3 & 1 & 4 \\ 1 & -4 & 2 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & -1 & 1 & 4 \\ -3 & -6 & 4 & 2 \\ 2 & 0 & -1 & 3 \\ 1 & -2 & 0 & -3 \end{bmatrix},$$

de forma que

$$A_{11} = \begin{bmatrix} 3 & -2 \\ 0 & 5 \end{bmatrix}; A_{12} = \begin{bmatrix} 3 & -1 \\ -1 & -3 \end{bmatrix}; A_{21} = \begin{bmatrix} -2 & 3 \\ 1 & -4 \end{bmatrix}; A_{22} = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$$

$$B_{11} = \begin{bmatrix} 2 & -1 \\ -3 & -6 \end{bmatrix}; B_{12} = \begin{bmatrix} 1 & 4 \\ 4 & 2 \end{bmatrix}; B_{21} = \begin{bmatrix} 2 & 0 \\ 1 & -2 \end{bmatrix}; B_{22} = \begin{bmatrix} -1 & 3 \\ 0 & -3 \end{bmatrix}$$

Encontremos $A + B$ y AB.

Solución

Aquí,

$$A+B=\begin{bmatrix} A_{11}+B_{11} & A_{12}+B_{12} \\ A_{21}+B_{21} & A_{22}+B_{22}\end{bmatrix}=\begin{bmatrix} 5 & -3 & 4 & 3 \\ -3 & -1 & 3 & -1 \\ 0 & 3 & 0 & 7 \\ 2 & -6 & 2 & 0\end{bmatrix}$$

y

$$AB=\begin{bmatrix} \begin{bmatrix} 12 & 9 \\ -15 & -30\end{bmatrix}+\begin{bmatrix} 5 & 2 \\ -5 & 6\end{bmatrix} & \begin{bmatrix} -5 & 8 \\ 20 & 10\end{bmatrix}+\begin{bmatrix} -3 & 12 \\ 1 & 6\end{bmatrix} \\ \begin{bmatrix} -13 & -16 \\ 14 & 23\end{bmatrix}+\begin{bmatrix} 6 & -8 \\ 7 & -6\end{bmatrix} & \begin{bmatrix} 10 & -2 \\ -15 & -4\end{bmatrix}+\begin{bmatrix} -1 & -9 \\ -2 & -3\end{bmatrix}\end{bmatrix}$$

$$=\begin{bmatrix} 17 & 11 & -8 & 20 \\ -20 & -24 & 21 & 16 \\ -7 & -24 & 9 & -11 \\ 21 & 17 & -17 & -7\end{bmatrix}$$

Ejemplo 25.

Consideremos las siguientes matrices A y B:

$$A=\begin{bmatrix} 2 & 1 & 0 & -2 \\ 5 & 7 & 4 & -1 \\ 3 & 1 & 6 & 4\end{bmatrix}, \qquad B=\begin{bmatrix} 2 & 7 \\ -1 & 1 \\ 0 & 3 \\ 4 & 2\end{bmatrix},$$

de forma que

$$A_{11}=\begin{bmatrix} 2 & 1 \\ 5 & 7\end{bmatrix}; A_{12}=\begin{bmatrix} 0 & -2 \\ 4 & -1\end{bmatrix}; A_{21}=\begin{bmatrix} 3 & 1\end{bmatrix}; A_{22}=\begin{bmatrix} 6 & 4\end{bmatrix}$$

$$B_{11}=\begin{bmatrix} 2 & 7 \\ -1 & 1\end{bmatrix}; \quad B_{21}=\begin{bmatrix} 0 & 3 \\ 4 & 2\end{bmatrix}$$

Encontremos AB.

Solución

$$A_{11}B_{11}=\begin{bmatrix} 3 & 15 \\ 3 & 42\end{bmatrix}; \quad A_{12}B_{21}=\begin{bmatrix} -8 & -4 \\ -4 & 10\end{bmatrix}$$

$$A_{21}B_{11}=\begin{bmatrix} 5 & 22\end{bmatrix}; \quad A_{22}B_{21}=\begin{bmatrix} 16 & 26\end{bmatrix}$$

$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} \\ A_{21}B_{11} + A_{22}B_{21} \end{bmatrix} = \begin{bmatrix} -5 & 11 \\ -1 & 52 \\ 21 & 48 \end{bmatrix}$$

Ejercicios 4

1) Halle la traspuesta de las siguientes matrices y determine si son (o no) simétricas:

a) $A = \begin{bmatrix} 7 & 9 & 6 & 5 \\ -2 & -3 & -1 & 0 \\ 4 & 8 & 7 & 1 \end{bmatrix}$ b) $A = \begin{bmatrix} 4 & 5 & -2 & 1 \\ 5 & 8 & 3 & 4 \\ -2 & 3 & 6 & -6 \\ 1 & 4 & -6 & -3 \end{bmatrix}$

c) $A = \begin{bmatrix} 1 & 2 & 3 & 5 & 7 \\ -10 & -12 & 4 & 8 & 14 \end{bmatrix}$ d) $A = \begin{bmatrix} 3 & 0 & 8 \\ 0 & 4 & 0 \\ 0 & 0 & 10 \end{bmatrix}$

e) $A = \begin{bmatrix} 3 & 1 \\ 8 & -2 \\ 6 & -3 \\ 1 & 9 \end{bmatrix}$ f) $A = \begin{bmatrix} 5 & 4 & 3 & 2 & 7 \\ 4 & 1 & -1 & 6 & 11 \\ 3 & -1 & 8 & 5 & 1 \\ 2 & 6 & 5 & -9 & 3 \\ 7 & 11 & 1 & 6 & 10 \end{bmatrix}$

2) Sean

$$A = \begin{bmatrix} 3 & 0 & 2 \\ 4 & 1 & 5 \end{bmatrix} \qquad B = \begin{bmatrix} 7 & 1 \\ 4 & 2 \\ 5 & 3 \end{bmatrix} \qquad C = \begin{bmatrix} 4 & 3 \\ 3 & 5 \end{bmatrix}$$

a) Calcule, si es posible, ABC, CBA, BCA, CB^TA^T y $C^TB^TA^T$.

b) Verifique que $(ABC)^T = C^TB^TA^T$.

3) Calcule, en cada caso, la traza de la matriz:

$$A = \begin{bmatrix} 4 & 6 & -3 \\ -2 & -3 & 0 \\ 7 & -8 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 2 & 0 & -1 & -3 \\ -2 & 4 & -5 & 2 \\ 0 & -6 & -2 & 1 \\ 9 & 0 & 5 & -4 \end{bmatrix}$$

$$C = \begin{bmatrix} 7 & -11 & -15 \\ 3 & 0 & 4 \\ -2 & 8 & 5 \end{bmatrix}$$

4) Dadas las matrices

$$A = \begin{bmatrix} -1 & 0 & -2 & -3 \\ 5 & 4 & -1 & 2 \\ -3 & -4 & 3 & 5 \\ 0 & 1 & 0 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & -2 & -4 & 0 \\ 0 & 5 & 1 & 3 \\ 4 & 0 & -1 & -5 \\ -1 & -3 & 0 & 2 \end{bmatrix}$$

encuentre $A + B$ y AB bajo las particiones:

a) Para A:

$$A_{11} = \begin{bmatrix} -1 & 0 \\ 5 & 4 \end{bmatrix} ; \; A_{12} = \begin{bmatrix} -2 & -3 \\ -1 & 2 \end{bmatrix}$$

$$A_{21} = \begin{bmatrix} -3 & -4 \\ 0 & 1 \end{bmatrix} ; \; A_{22} = \begin{bmatrix} 3 & 5 \\ 0 & 2 \end{bmatrix}$$

b) Para B:

$$B_{11} = \begin{bmatrix} 3 & -2 \\ 0 & 5 \end{bmatrix} ; \; B_{12} = \begin{bmatrix} -4 & 0 \\ 1 & 3 \end{bmatrix}$$

$$B_{21} = \begin{bmatrix} 4 & 0 \\ -1 & -3 \end{bmatrix} ; \; B_{22} = \begin{bmatrix} -1 & -5 \\ 0 & 2 \end{bmatrix}$$

5. Determinante de una matriz cuadrada

Un sistema de dos ecuaciones lineales con dos incógnitas de la forma

$$\begin{aligned} a_{11}x + a_{12}y &= b_1 \\ a_{21}x + a_{22}y &= b_2 \end{aligned}$$

se puede resolver multiplicando la primera ecuación por a_{22} y la segunda por $-a_{12}$, y luego sumándolas. Así obtenemos

$$(\,a_{11}a_{22} - a_{21}a_{12}\,)x = b_1a_{22} - b_2a_{12}$$

y, de manera similar, multiplicando la primera ecuación por $-a_{21}$ y la segunda por a_{11}, y sumando, obtenemos

$$(\,a_{11}a_{22} - a_{21}a_{12}\,)y = a_{11}b_2 - a_{21}b_1$$

De estas ecuaciones es fácil determinar x y y si la expresión $a_{11}a_{22} - a_{21}a_{12}$ es diferente de cero. Precisamente a esta expresión se le llama el *determinante* de la matriz de coeficientes del sistema (término acuñado por Karl. F. Gauss en 1801), y su conveniencia se verá claramente en los siguientes desarrollos teóricos.

a. Determinantes 2×2

Comenzamos entonces definiendo los determinantes más simples posibles: los de las matrices cuadradas 2×2.

Definición 15. (Determinante de una matriz 2×2)
Para una matriz 2×2

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

su *determinante* está dado por el valor $|A| = a_{11}a_{22} - a_{21}a_{12}$, que también se denota con el símbolo $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ (barras en lugar de paréntesis abarcando los números de la matriz), o con $\det A$.

Nota 3.
La notación de barras anterior fue introducida por Augustin L. Cauchy en 1826, aunque algunos (equivocadamente) la asocian con Cayley (1841).

Ejemplo 26.
Utilizando la fórmula de la definición 15 obtenemos que

$$\begin{vmatrix} 1 & 3 \\ 4 & 5 \end{vmatrix} = (1)(5) - (4)(3) = -7 \quad \text{y} \quad \begin{vmatrix} 1 & 7 \\ 2 & 14 \end{vmatrix} = (1)(14) - (2)(7) = 0$$

$$\begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = (2)(2) - (0)(0) = 4 \quad \text{y} \quad \begin{vmatrix} -2 & -3 \\ 3 & 1 \end{vmatrix} = (-2)(1) - (-3)(3) = 7 \quad \blacktriangle$$

Una profunda (y, por tanto, simple) conexión del concepto de determinante con nociones geométricas nos comienza a mostrar que nuestro concepto aquí está muy relacionado con las figuras *conformadas por rectas y planos*:

Ejemplo 27. (Área del paralelogramo (Cayley (1841)))
Consideremos el paralelogramo "generado" por los puntos (a_{11}, a_{21}) y (a_{12}, a_{22}) (figura 1a). Veamos que su área es precisamente igual al valor absoluto del determinante de la matriz

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

es decir, el área del paralelogramo es $|a_{11}a_{22} - a_{12}a_{21}|$.

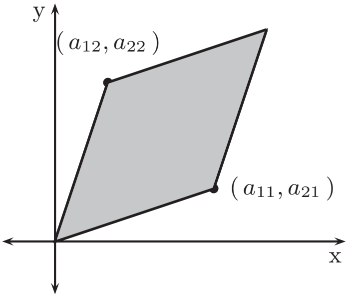


Figura 1a

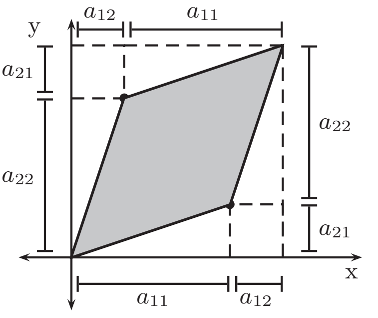


Figura 1b

Solución

A partir de la figura 1b) podemos observar que el área del paralelogramo es igual a

$$(a_{11} + a_{12})(a_{21} + a_{22}) - 2a_{12}a_{21} - 2\frac{a_{12}a_{22}}{2} - 2\frac{a_{11}a_{21}}{2} = a_{11}a_{22} - a_{12}a_{21}$$

Así, el área del paralelogramo es el determinante de A. Observemos que si un punto fuera un múltiplo escalar de otro, los dos puntos estarían sobre la misma línea. En este caso, el paralelogramo colapsaría en una línea y tendría un área igual a cero.

Ejemplo 28.

De acuerdo con el ejemplo 27 anterior, el área del paralelogramo generado por los puntos $(4, 1)$ y $(7, 3)$ es

$$\begin{vmatrix} 4 & 7 \\ 1 & 3 \end{vmatrix} = 5$$

En la lección 6 (*Transformaciones lineales*) entenderemos el trasfondo de esta conexión entre determinantes y geometría.

b. Determinantes 3×3

De manera similar a lo ya estudiado para sistemas 2×2, podemos buscar soluciones a un sistema 3×3 en el cual encontraremos también el término que llamaremos *determinante de una matriz* 3×3.

Definición 16. (Determinante de una matriz 3×3)

El *determinante* de la matriz 3×3

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

se puede calcular mediante la siguiente fórmula:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Ejemplo 29.
Calculemos los determinantes de las matrices

$$\text{a)}\quad A = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & 5 \\ 4 & 5 & 9 \end{bmatrix} \qquad \text{b)}\quad B = \begin{bmatrix} 3 & -1 & 2 \\ 1 & 2 & 4 \\ 6 & 3 & 5 \end{bmatrix}$$

Solución

a)

$$\begin{aligned} \begin{vmatrix} 1 & 0 & -2 \\ 3 & 1 & 5 \\ 4 & 5 & 9 \end{vmatrix} &= 1\begin{vmatrix} 1 & 5 \\ 5 & 9 \end{vmatrix} - 0\begin{vmatrix} 3 & 5 \\ 4 & 9 \end{vmatrix} + (-2)\begin{vmatrix} 3 & 1 \\ 4 & 5 \end{vmatrix} \\ &= 1[(1)(9) - (5)(5)] + (-2)[(3)(5) - (1)(4)] \\ &= -38 \end{aligned}$$

b)

$$\begin{aligned} \begin{vmatrix} 3 & -1 & 2 \\ 1 & 2 & 4 \\ 6 & 3 & 5 \end{vmatrix} &= 3\begin{vmatrix} 2 & 4 \\ 3 & 5 \end{vmatrix} - (-1)\begin{vmatrix} 1 & 4 \\ 6 & 5 \end{vmatrix} + 2\begin{vmatrix} 1 & 2 \\ 6 & 3 \end{vmatrix} \\ &= 3[(2)(5) - (4)(3)] + 1[(1)(5) - (4)(6)] \\ &\quad + 2[(1)(3) - (2)(6)] \\ &= -43 \quad \blacktriangle \end{aligned}$$

Y nuevamente encontramos una conexión entre determinantes y geometría, ahora en el caso de los determinantes 3×3. Es la siguiente:

Ejemplo 30. (Longitud de un segmento y área de un triángulo (Cayley (1841)))

a) Ya sabemos que la longitud de un segmento que va desde a_{21} hasta a_{11} con $a_{11} > a_{21}$ en la recta real, está medida por $a_{11} - a_{21}$; pero obsérvese que esto se puede escribir así:

$$a_{11} - a_{21} = \begin{vmatrix} a_{11} & 1 \\ a_{21} & 1 \end{vmatrix}$$

b) Consideremos ahora el triángulo en el plano cuyos vértices son $P_1(a_{11}, a_{12})$, $P_2(a_{21}, a_{22})$ y $P_3(a_{31}, a_{32})$(figura 2).

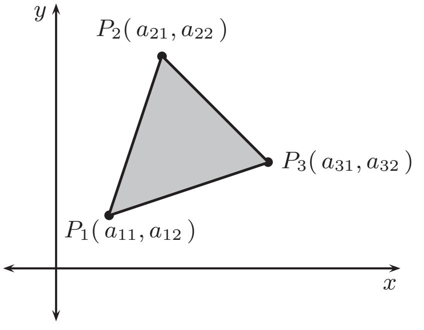


Figura 2. Área de un triángulo

Es fácil mostrar (utilizando la fórmula y el procedimiento del ejemplo 27) que el área de este triángulo es igual al valor absoluto del determinante de la matriz

$$A = \frac{1}{2}\begin{bmatrix} a_{11} & a_{12} & 1 \\ a_{21} & a_{22} & 1 \\ a_{31} & a_{32} & 1 \end{bmatrix}$$

Solución

Observemos que

$$\begin{aligned} \frac{1}{2}\begin{vmatrix} a_{11} & a_{12} & 1 \\ a_{21} & a_{22} & 1 \\ a_{31} & a_{32} & 1 \end{vmatrix} &= \frac{1}{2}\left[(a_{32} - a_{12})(a_{21} - a_{11}) - (a_{22} - a_{12})(a_{31} - a_{11})\right] \\ &= \frac{1}{2}\begin{vmatrix} a_{21} - a_{11} & a_{31} - a_{11} \\ a_{22} - a_{12} & a_{32} - a_{12} \end{vmatrix} \end{aligned}$$

El área del triángulo cuyos vértices son (a_{11}, a_{12}), (a_{21}, a_{22}) y (a_{31}, a_{32}) es igual al área del triángulo cuyos vértices son $(0, 0)$, $(a_{21} - a_{11}, a_{22} - a_{12})$ y $(a_{31} - a_{11}, a_{32} - a_{12})$. Pero utilizando el resultado del ejemplo 27, sabemos que el área del paralelogramo generado por los puntos $(a_{21} - a_{11}, a_{22} - a_{12})$ y $(a_{31} - a_{11}, a_{32} - a_{12})$ es el determinante de la matriz

$$\begin{bmatrix} a_{21} - a_{11} & a_{31} - a_{11} \\ a_{22} - a_{12} & a_{32} - a_{12} \end{bmatrix}$$

Como el área del triángulo en cuestión es la mitad del área de este paralelogramo (figura 3), se tiene el resultado que queríamos mostrar.

Ejemplo 31.

Calculemos el área del triángulo cuyos vértices son $(1, -4)$, $(4, 8)$ y $(1, 1)$.

Solución

De acuerdo con el ejemplo 30, el área del triángulo formado por los puntos $(1,-4)$, $(4,8)$ y $(1,1)$ es

$$\frac{1}{2}\begin{vmatrix} 1 & -4 & 1 \\ 4 & 8 & 1 \\ 1 & 1 & 1 \end{vmatrix} = \frac{15}{2}$$

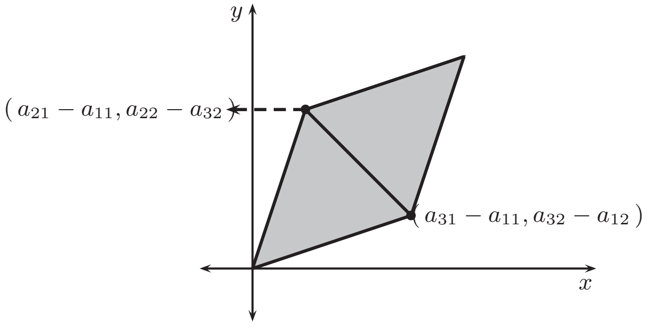


Figura 3

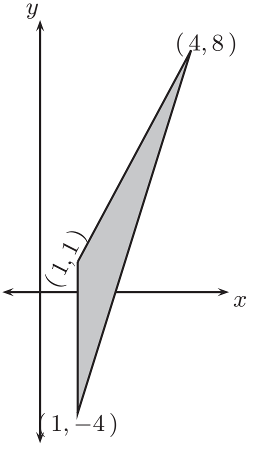


Figura 4

Ejemplo 32. (Volumen del paralelepípedo como un determinante)
De manera similar a lo estudiado con el área del paralelogramo en el ejemplo 27, se puede mostrar que, geométricamente, el valor absoluto del determinante

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

es el *volumen del paralelepípedo* en el espacio tridimensional, formado por los puntos $P_1(a_{11}, a_{12}, a_{13})$, $P_2(a_{21}, a_{22}, a_{23})$ y $P_3(a_{31}, a_{32}, a_{33})$ (figura 5). Mostrar esto queda como ejercicio para el lector.

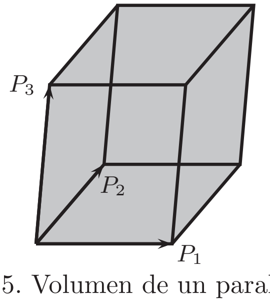


Figura 5. Volumen de un paralelepípedo

Ejemplo 33.
Calculemos el volumen del paralelepípedo formado por los puntos $P_1(2,0,1)$, $P_2(1,3,1)$, $P_3(1,2,1)$.

Solución

El volumen del paralelepípedo es

$$V = \begin{vmatrix} 2 & 0 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 1 \end{vmatrix} = (2)(1) + (1)(-1) = 1$$

c. Determinantes $n \times n$

El siguiente paso en la generalización del concepto de determinante es la definición del determinante de una matriz $n \times n$, y esto lo efectuaremos de manera recursiva; es decir, para calcular el determinante de una matriz de tamaño $n \times n$, supondremos que ya sabemos cómo calcular el determinante de una matriz de tamaño $(n-1) \times (n-1)$. Así, puesto que ya sabemos cómo calcular el determinante de matrices 2×2 y 3×3, entonces podemos calcular el determinante de matrices 4×4 y, en consecuencia, podemos también calcular el de matrices 5×5, etc.

Definición 17. (Determinante de una matriz $n \times n$)
Sea A una matriz $n \times n$ cualquiera, y sea A_{1j} la matriz obtenida de A eliminando la primera fila y la j-ésima columna. El *determinante* de A está dado por la fórmula

$$|A| = a_{11}\,|A_{11}| - a_{12}\,|A_{12}| + a_{13}\,|A_{13}| + \cdots + (-1)^{n-1}a_{1n}\,|A_{1n}|$$

Esta expresión se denomina *expansión por cofactores*. Aquí los cofactores son A_{11}, A_{12}, A_{13},...,A_{1n}. Obsérvese también la alternancia en el signo de los sumandos.

Ejemplo 34.
Calculemos el siguiente determinante:

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \\ 3 & 2 & 1 & 4 \\ 2 & 1 & 4 & 3 \end{vmatrix}$$

Solución
Utilizando la expansión por cofactores tenemos que este determinante es

$$1\begin{vmatrix} 3 & 2 & 1 \\ 2 & 1 & 4 \\ 1 & 4 & 3 \end{vmatrix} - 2\begin{vmatrix} 4 & 2 & 1 \\ 3 & 1 & 4 \\ 2 & 4 & 3 \end{vmatrix} + 3\begin{vmatrix} 4 & 3 & 1 \\ 3 & 2 & 4 \\ 2 & 1 & 3 \end{vmatrix} - 4\begin{vmatrix} 4 & 3 & 2 \\ 3 & 2 & 1 \\ 2 & 1 & 4 \end{vmatrix}$$

y tenemos que calcular cada uno de estos determinantes 3×3. Pero el primer determinante es

$$\begin{aligned}\begin{vmatrix} 3 & 2 & 1 \\ 2 & 1 & 4 \\ 1 & 4 & 3 \end{vmatrix} &= 3\begin{vmatrix} 1 & 4 \\ 4 & 3 \end{vmatrix} - 2\begin{vmatrix} 2 & 4 \\ 1 & 3 \end{vmatrix} + 1\begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} \\ &= 3[(3)(1) - (4)(4)] - 2[(2)(3) - (4)(1)] \\ &\quad + 1[(2)(4) - (1)(1)] \\ &= -39 - 4 + 7 = -36\end{aligned}$$

Y, de manera similar, los valores para los otros tres determinantes son -44, 4, y -4, respectivamente. Por lo tanto, el determinante pedido es

$$1(-36) - 2(-44) + 3(4) - 4(-4) = 80 \quad \blacktriangle$$

Ahora: en la fórmula de expansión por cofactores de la definición inmediatamente anterior pareciera que la primera fila de la matriz juega un papel especial en el cálculo del determinante. Sin embargo, el siguiente teorema afirma que esto no es cierto, pues el determinante de una matriz puede calcularse expandiendo a través de *cualquier fila o cualquier columna*. Veamos este teorema, aunque su prueba no será presentada aquí.

Teorema 7. ***(Cálculo del determinante por cofactores (Laplace (1772)))***
Sean A una matriz $n \times n$ y A_{ij} $(i, j = 1, 2, \ldots, n)$ la matriz obtenida de A eliminando la i-ésima fila y la j-ésima columna. Entonces

a) Para cada fila i,

$$|A| = (-1)^{i+1}(a_{i1}|A_{i1}| - a_{i2}|A_{i2}| + a_{i3}|A_{i3}| + \cdots + (-1)^{n-1}a_{in}|A_{in}|)$$

b) Para cada columna j,

$$|A| = (-1)^{j+1}(a_{1j}|A_{1j}| - a_{2j}|A_{2j}| + a_{3j}|A_{3j}| + \cdots + (-1)^{n-1}a_{nj}|A_{nj}|)$$

A cada uno de los términos de arriba de la forma A_{ij} se le denomina un cofactor de A. Al igual que en la definición 17, a esta expresión también se le conoce como expansión por cofactores.

Ejemplo 35.
Calculemos de nuevo el determinante del ejemplo 29a) expandiendo, por ejemplo, por la segunda fila. Observemos que, en este caso, la expansión va multiplicada por el factor $(-1)^{2+1} = -1$. Así,

$$\begin{vmatrix} 1 & 0 & -2 \\ 3 & 1 & 5 \\ 4 & 5 & 9 \end{vmatrix} = -3\begin{vmatrix} 0 & -2 \\ 5 & 9 \end{vmatrix} + 1\begin{vmatrix} 1 & -2 \\ 4 & 9 \end{vmatrix} - 5\begin{vmatrix} 1 & 0 \\ 4 & 5 \end{vmatrix}$$

$$= -30 + 17 - 25 = -38$$

Ejemplo 36.
Calculemos de nuevo el determinante del ejemplo 29a) expandiendo ahora por la segunda columna. Como en el ejemplo anterior, la expansión va multiplicada por el factor $(-1)^{2+1} = -1$. Así,

$$\begin{vmatrix} 1 & 0 & -2 \\ 3 & 1 & 5 \\ 4 & 5 & 9 \end{vmatrix} = -0\begin{vmatrix} 3 & 5 \\ 4 & 9 \end{vmatrix} + 1\begin{vmatrix} 1 & -2 \\ 4 & 9 \end{vmatrix} - 5\begin{vmatrix} 1 & -2 \\ 3 & 5 \end{vmatrix} = 0 + 17 - 55 = -38$$

Ejemplo 37.
Al tratar de calcular el determinante de la matriz

$$A = \begin{bmatrix} 1 & -1 & 0 & 2 \\ 0 & -2 & 0 & 1 \\ -3 & 0 & \frac{1}{2} & 4 \\ 0 & 3 & 0 & -1 \end{bmatrix}$$

inmediatamente podemos notar que si lo calculamos expandiendo por la tercera columna, ahorramos gran cantidad de cálculos, ya que la mayor parte de los términos en esta columna son ceros. Luego,

$$|A| = (-1)^{(3+1)}\frac{1}{2}\begin{vmatrix} 1 & -1 & 2 \\ 0 & -2 & 1 \\ 0 & 3 & -1 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} -2 & 1 \\ 3 & -1 \end{vmatrix} = -\frac{1}{2}$$

Ejemplo 38.
Calculemos el determinante de la matriz

$$A = \begin{bmatrix} 1 & -3 & 0 & -2 \\ 3 & -12 & -2 & -6 \\ -2 & 10 & 2 & 5 \\ -1 & 6 & 1 & 3 \end{bmatrix}$$

Solución

Expandiendo por la primera fila se tiene que

$$
\begin{aligned}
|\,A\,| &= 1\begin{vmatrix} -12 & -2 & -6 \\ 10 & 2 & 5 \\ 6 & 1 & 3 \end{vmatrix} + 3\begin{vmatrix} 3 & -2 & -6 \\ -2 & 2 & 5 \\ -1 & 1 & 3 \end{vmatrix} + 2\begin{vmatrix} 3 & -12 & -2 \\ -2 & 10 & 2 \\ -1 & 6 & 1 \end{vmatrix} \\
&= (\,1\,)(\,0\,) + (\,3\,)(\,1\,) + (\,2\,)(\,-2\,) = -1
\end{aligned}
$$

Nota 4.

Para evitar errores posibles, resaltemos una vez más los signos que acompañan a los cofactores en una matriz 3×3:

$$
\begin{bmatrix} (+) & (-) & (+) \\ (-) & (+) & (-) \\ (+) & (-) & (+) \end{bmatrix}
$$

¿Cuáles serán estos signos si la matriz es 4×4? ¿5×5?

Ejemplo 39. (Volumen de un tetraedro (Cayley (1841)))

De manera similar a lo obtenido en el ejemplo 30, dados cuatro puntos en el espacio cartesiano, $P_1(a_{11}, a_{12}, a_{13})$, $P_2(a_{21}, a_{22}, a_{23})$, $P_3(a_{31}, a_{32}, a_{33})$, y $P_4(a_{41}, a_{42}, a_{43})$, el volumen del tetraedro con vértices en P_1, P_2, P_3, y P_4, (Volumen 0 (Fundamentos)) está dado por el valor absoluto del determinante

$$
\frac{1}{6}\begin{vmatrix} a_{11} & a_{12} & a_{13} & 1 \\ a_{21} & a_{22} & a_{23} & 1 \\ a_{31} & a_{32} & a_{33} & 1 \\ a_{41} & a_{42} & a_{43} & 1 \end{vmatrix}
$$

Esta afirmación puede comprobarse expandiendo el determinante y se deja como ejercicio (laborioso) para el lector (recuerde que el área de un tetraedro es un tercio del área de la base multiplicada por la altura). Obsérvese cómo las fórmulas de longitud y área del triángulo (ejemplo 30) y esta del volumen del tetraedro (ejemplo 39), comparten una notable similitud cuando se expresan en términos de determinantes. Más adelante, el lector podría aventurar una hipótesis de por qué esto es así, una vez conozca las propiedades de los determinantes.

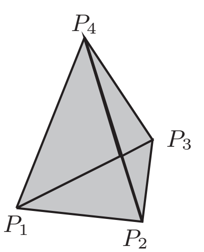


Figura 6. Volumen de un tetraedro

Ejemplo 40.
Calculemos el volumen del tetraedro formado por los cuatro puntos del espacio $P_1(2,2,2)$, $P_2(4,1,3)$, $P_3(5,3,4)$ y $P_4(3,5,1)$.

Solución
El volumen de este tetraedro es el valor absoluto del determinante

$$\frac{1}{6}\begin{vmatrix} 2 & 2 & 2 & 1 \\ 4 & 1 & 3 & 1 \\ 5 & 3 & 4 & 1 \\ 3 & 5 & 1 & 1 \end{vmatrix} = \frac{1}{6}(11) = \frac{11}{6}$$

Ejercicios 5

1) Calcule los determinantes de las siguientes matrices:

a) $A = \begin{bmatrix} 4 & 2 & 1 \\ 0 & 3 & 5 \\ 8 & -1 & 2 \end{bmatrix}$ b) $A = \begin{bmatrix} 0 & -2 & 1 \\ 1 & 0 & -2 \\ 6 & 3 & 6 \end{bmatrix}$

c) $A = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 6 & -2 & 5 & 10 \\ -3 & 0 & 1 & 0 \\ 4 & 0 & 0 & 7 \end{bmatrix}$ d) $A = \begin{bmatrix} 6 & 4 & 2 & 2 \\ 6 & 8 & -4 & 2 \\ 0 & 3 & 0 & 0 \\ 3 & 9 & 2 & 1 \end{bmatrix}$

e) $A = \begin{bmatrix} 4 & 7 \\ 5 & 9 \end{bmatrix}$ f) $A = \begin{bmatrix} -3 & 10 \\ -5 & 18 \end{bmatrix}$

2) Demuestre que

a) $\det(I_n) = 1$ para todo n b) $\det(0) = 0$

3) Muestre que

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0$$

es la ecuación de la recta que pasa por los puntos $P_1(x_1, y_1)$ y $P_2(x_2, y_2)$. Halle la recta que pasa por los puntos $(2,-3)$ y $(3,4)$.

4) Muestre que

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0$$

es la ecuación del plano que pasa por los puntos $P_1(x_1, y_1, z_1)$, $P_2(x_2, y_2, z_2)$ y $P_3(x_3, y_3, z_3)$. Encuentre el plano que pasa por los puntos (1, –5,2), (2,5,–1) y (6,1,4).

5) Muestre que la circunferencia que pasa por los puntos $P_1(x_1, y_1)$, $P_2(x_2, y_2)$, $P_3(x_3, y_3)$ en el plano, está dada por la ecuación

$$\begin{vmatrix} x^2+y^2 & x & y & 1 \\ x_1^2+y_1^2 & x_1 & y_1 & 1 \\ x_2^2+y_2^2 & x_2 & y_2 & 1 \\ x_3^2+y_3^2 & x_3 & y_3 & 1 \end{vmatrix} = 0$$

6) Encuentre el área del paralelogramo formado por los siguientes puntos:

a) $(3, 9)$ y $(1, 7)$ b) $(5, 6)$ y $(4, 5)$

7) Encuentre el área del triángulo cuyos vértices son:

a) $(-2, 1), (3, 7)$ y $(6, -3)$ b) $(-1, -4), (4, 6)$ y $(2, 2)$

8) Encuentre el volumen del tetraedro formado por los cuatro puntos del espacio $P_1(2, 2, 2)$, $P_2(-1, 2, 3)$, $P_3(4, -3, 7)$ y $P_4(0, 5, 8)$.

6. Propiedades de los determinantes

En esta sección examinaremos las propiedades más importantes que satisface el concepto de determinante de una matriz. Esto nos permitirá comprender sus propiedades lineales y, así, describir otras formas de calcularlo que podrían facilitar su evaluación en problemas concretos.

Teorema 8.
Para cualquier matriz A de orden $n \times n$, $\det A = \det A^T$.

Demostración
Es consecuencia del teorema 7, pues cualquier determinante se puede evaluar a través de filas o de columnas. ■

Ejemplo 41.
Calculemos el determinante de las matrices traspuestas de las matrices de los ejemplos 29a) y 34.

Solución

a) $A^T = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & 5 \\ 4 & 5 & 9 \end{bmatrix}^T = \begin{bmatrix} 1 & 3 & 4 \\ 0 & 1 & 5 \\ -2 & 5 & 9 \end{bmatrix}$

El determinante de A^T es

$$1\begin{vmatrix} 1 & 5 \\ 5 & 9 \end{vmatrix} - 3\begin{vmatrix} 0 & 5 \\ -2 & 9 \end{vmatrix} + 4\begin{vmatrix} 0 & 1 \\ -2 & 5 \end{vmatrix} = (1)(-16) - (3)(10) + (4)(2)$$
$$= -38 = \det A$$

b) $B^T = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \\ 3 & 2 & 1 & 4 \\ 2 & 1 & 4 & 3 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 & 3 & 2 \\ 2 & 3 & 2 & 1 \\ 3 & 2 & 1 & 4 \\ 4 & 1 & 4 & 3 \end{bmatrix}$; y el determinante de B^T es

$$1\begin{vmatrix} 3 & 2 & 1 \\ 2 & 1 & 4 \\ 1 & 4 & 3 \end{vmatrix} - 4\begin{vmatrix} 2 & 2 & 1 \\ 3 & 1 & 4 \\ 4 & 4 & 3 \end{vmatrix} + 3\begin{vmatrix} 2 & 3 & 1 \\ 3 & 2 & 4 \\ 4 & 1 & 3 \end{vmatrix} - 2\begin{vmatrix} 2 & 3 & 2 \\ 3 & 2 & 1 \\ 4 & 1 & 4 \end{vmatrix}$$
$$= -36 + 16 + 60 + 40$$
$$= 80 = \det B$$

Teorema 9. ***(Propiedad escalar del determinante)***
Si una matriz B se obtiene de una matriz A multiplicando cada elemento de una fila cualquiera de A por un escalar k, entonces $\det B = k \det A$.

Demostración
De acuerdo con el teorema 7 anterior, si j es la columna de A multiplicada por k, entonces

$$\begin{aligned} |B| &= (-1)^{j+1}(ka_{1j}|A_{1j}| - ka_{2j}|A_{2j}| + ka_{3j}|A_{3j}| + \cdots + (-1)^{n-1}ka_{nj}|A_{nj}|) \\ &= k(-1)^{j+1}(a_{1j}|A_{1j}| - a_{2j}|A_{2j}| + a_{3j}|A_{3j}| + \cdots + (-1)^{n-1}a_{nj}|A_{nj}|) \\ &= k|A| \quad \blacksquare \end{aligned}$$

Ejemplo 42.
Sean

$$A = \begin{bmatrix} 2 & 5 \\ 3 & -1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 6 & 15 \\ 3 & -1 \end{bmatrix}$$

Notemos que B se obtiene de A multiplicando la primera fila de ésta por 3. Los determinantes de estas matrices son:

$$\det A = 2(-1) - 5(3) = -17$$
$$\det B = 6(-1) - 15(3) = -51 = 3\det A$$

Ejemplo 43.
Calculemos el determinante de la matriz

$$\begin{bmatrix} -6 & 8 & 2 \\ 15 & 20 & 5 \\ 3 & 4 & -1 \end{bmatrix}$$

Solución

Utilizando las propiedades del determinante se tiene que

$$\begin{aligned}
\begin{vmatrix} -6 & 8 & 2 \\ 15 & 20 & 5 \\ 3 & 4 & -1 \end{vmatrix} = 2 \cdot \begin{vmatrix} -3 & 4 & 1 \\ 15 & 20 & 5 \\ 3 & 4 & -1 \end{vmatrix} &= 2 \cdot 5 \cdot \begin{vmatrix} -3 & 4 & 1 \\ 3 & 4 & 1 \\ 3 & 4 & -1 \end{vmatrix} \\
&= 2 \cdot 5 \cdot 3 \begin{vmatrix} -1 & 4 & 1 \\ 1 & 4 & 1 \\ 1 & 4 & -1 \end{vmatrix} = 2 \cdot 5 \cdot 3 \cdot 4 \begin{vmatrix} -1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \end{vmatrix} \\
&= 480 \quad \blacktriangle
\end{aligned}$$

Como una consecuencia de la propiedad escalar del determinante se tiene el siguiente resultado:

Corolario 1

Si A es una matriz cuadrada de tamaño $n \times n$, y k un escalar, entonces

$$\det(\, kA \,) = k^n \det A$$

Teorema 10. ***(Propiedad aditiva del determinante)***
Denotemos el vector fila i de la matriz A de orden $n \times n$ por A_i, donde $i = 1, 2, \ldots, n$ y supongamos, en particular, que $A_i = U + V$, donde U y V son matrices $1 \times n$. Entonces,

$$\det A = \det \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ U + V \\ \vdots \\ A_n \end{bmatrix} = \det \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ U \\ \vdots \\ A_n \end{bmatrix} + \det \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ V \\ \vdots \\ A_n \end{bmatrix}$$

Esta misma propiedad se cumple si una fila cualquiera A_i es la suma de matrices $n \times 1$, U y V.

Demostración

Sea

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ b_{i1} + c_{i1} & \cdots & b_{in} + c_{in} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

Entonces, expandiendo por la fila i, se tiene que

$$\begin{aligned} \det A &= (-1)^{i+1}[(b_{i1}+c_{i1})|A_{i1}| - (b_{i2}+c_{i2})|A_{i2}| + \cdots + (b_{in}+c_{in})|A_{in}|] \\ &= (-1)^{i+1}[b_{i1}|A_{i1}| - b_{i2}|A_{i2}| + \cdots + b_{in}|A_{in}|] + (-1)^{i+1}[c_{i1}|A_{i1}| - \\ &\quad c_{i2}|A_{i2}| + \cdots + c_{in}|A_{in}|] \\ &= \det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ b_{i1} & \cdots & b_{in} \\ \vdots & \cdots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} + \det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ c_{i1} & \cdots & c_{in} \\ \vdots & \cdots & \vdots \\ c_{m1} & \cdots & c_{mn} \end{bmatrix} \quad \blacksquare \end{aligned}$$

Ejemplo 44.

a) Partiendo la tercera fila en $(2,0,2) = (2,0,0) + (0,0,2)$, se tiene que

$$\begin{vmatrix} 1 & 1 & 3 \\ 2 & -1 & 4 \\ 2 & 0 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 3 \\ 2 & -1 & 4 \\ 2 & 0 & 0 \end{vmatrix} + \begin{vmatrix} 1 & 1 & 3 \\ 2 & -1 & 4 \\ 0 & 0 & 2 \end{vmatrix} = 14 + (-6) = 8$$

b) Partiendo la segunda fila en $(0,1,6) = (0,0,5) + (0,1,1)$, se tiene que

$$\begin{vmatrix} 1 & 3 & 4 \\ 0 & 1 & 6 \\ -3 & 4 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 4 \\ 0 & 0 & 5 \\ -3 & 4 & 9 \end{vmatrix} + \begin{vmatrix} 1 & 3 & 4 \\ 0 & 1 & 1 \\ -3 & 4 & 9 \end{vmatrix} = -65 + 8 = -57 \quad \blacktriangle$$

Nos preguntamos ahora cómo varía el determinante de una matriz si intercambiamos dos filas cualesquiera de la matriz. La siguiente propiedad responde a esta pregunta:

Teorema 11. ***(Propiedad de intercambio de filas del determinante)*** *Si una matriz B de orden $n \times n$ se obtiene de una matriz A intercambiando dos filas cualesquiera de A, entonces*

$$\det B = -\det A$$

En particular, si dos filas de la matriz A son iguales, el determinante de A es cero.

Demostración

Consideremos inicialmente que se intercambian dos filas adyacentes: la i-ésima y la $(i+1)$-ésima. Las matrices A y B son

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \quad \text{y} \quad B=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

Desarrollando el determinante de A por la i-ésima fila y el determinante de B por la $(i+1)$-ésima fila se tiene que

$$\det A = (-1)^{i+1}(a_{i1}\,|A_{i1}| - a_{i2}\,|A_{i2}| + a_{i3}\,|A_{i3}| + \cdots + (-1)^{n-1}a_{in}\,|A_{in}|)$$

$$\det B = (-1)^{i+2}(a_{i1}\,|B_{i+1,1}| - a_{i2}\,|B_{i+1,2}| + a_{i3}\,|B_{i+1,3}| + \cdots + (-1)^{n-1}a_{in}\,|B_{i+1,n}|)$$

Notemos ahora que si eliminamos la $(i+1)$-ésima fila y la j-ésima columna de B, obtenemos A_{ij}; es decir, $B_{i+1,j} = A_{ij}$. Por tanto,

$$\begin{aligned} \det B &= (-1)^{i+2}(a_{i1}\,|A_{i1}| - a_{i2}\,|A_{i2}| + a_{i3}\,|A_{i3}| + \cdots + (-1)^{n-1}a_{in}\,|A_{in}|) \\ &= -1(-1)^{i+1}(a_{i1}\,|A_{i1}| - a_{i2}\,|A_{i2}| + a_{i3}\,|A_{i3}| + \cdots + (-1)^{n-1}a_{in}\,|A_{in}|) \\ &= -\det A \end{aligned}$$

Supongamos ahora que $i < k$ y que se intercambian la i-ésima y la k-ésima fila. Esto se realiza intercambiando varias veces filas adyacentes, y observemos que es necesario realizar $k-i$ intercambios adyacentes para mover la k-ésima fila a la i-ésima fila. Se requiere, además, realizar $k-i-1$ intercambios adyacentes para mover la i-ésima fila a la k-ésima fila. Así, el número total de intercambios adyacentes es $k-i+k-i-1 = 2(k-i)-1$, el cual es un número impar. Por tanto, $\det B = -\det A$. ■

Ejemplo 45.

Veamos cómo se relaciona el determinante de la matriz A con el determinante de la matriz B, donde A y B son

$$A=\begin{bmatrix} 1 & 6 & 3 \\ 1 & 1 & 1 \\ 3 & 4 & 5 \end{bmatrix}, \qquad B=\begin{bmatrix} 3 & 4 & 5 \\ 1 & 6 & 3 \\ 4 & 4 & 4 \end{bmatrix}$$

Solución

Aplicando la propiedad escalar del determinante y la propiedad de intercambio de filas sobre la matriz B se concluye que

$$\det B = 4\begin{vmatrix} 3 & 4 & 5 \\ 1 & 6 & 3 \\ 1 & 1 & 1 \end{vmatrix} = -4\begin{vmatrix} 1 & 6 & 3 \\ 3 & 4 & 5 \\ 1 & 1 & 1 \end{vmatrix} = 4\begin{vmatrix} 1 & 6 & 3 \\ 1 & 1 & 1 \\ 3 & 4 & 5 \end{vmatrix} = 4 \ \det A \qquad \blacktriangle$$

El siguiente teorema relaciona el determinante de un producto de matrices con el determinante de cada una de las matrices que forman el producto:

Teorema 12. ***(Determinante del producto)***
Para todo par de matrices A y B de orden $n \times n$ se tiene que

$$\det(\,AB\,) = \det A \ \det B$$

Demostración

Sólo demostraremos el caso $n = 2$. El caso $n > 2$ es similar, aunque más demandante en términos de cálculos algebraicos.

Sean $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ y $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$; entonces

$$AB = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

Así,

$$\begin{aligned} \det A \det B &= (a_{11}a_{22} - a_{12}a_{21})(b_{11}b_{22} - b_{12}b_{21}) \\ &= a_{11}a_{22}b_{11}b_{22} - a_{11}a_{22}b_{12}b_{21} - a_{12}a_{21}b_{11}b_{22} + a_{12}a_{21}b_{12}b_{21} \end{aligned}$$

$$\begin{aligned} \text{y} \quad \det(AB) &= (a_{11}b_{11} + a_{12}b_{21})(a_{21}b_{12} + a_{22}b_{22}) - \\ &\quad\ (a_{11}b_{12} + a_{12}b_{22})(a_{21}b_{11} + a_{22}b_{21}) \\ &= a_{11}a_{22}b_{11}b_{22} - a_{11}a_{22}b_{12}b_{21} - a_{12}a_{21}b_{11}b_{22} + a_{12}a_{21}b_{12}b_{21} \end{aligned}$$

Luego

$$\det(AB) = \det A \det B \qquad \blacksquare$$

Ejemplo 46.
Sean

a) $A = \begin{bmatrix} 1 & -4 \\ -3 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 5 & 7 \\ -3 & 5 \end{bmatrix}$

b) $A = \begin{bmatrix} 1 & 4 & -3 \\ 5 & 1 & 8 \\ 0 & 7 & 6 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 3 & 6 \\ -1 & 2 & 5 \\ 4 & 8 & 7 \end{bmatrix}$

Veamos que, en cada caso, $\det(AB) = \det A \det B$.

Solución

a) El determinante de A es igual a -10 y el determinante de B es igual a 46. Luego, $\det A \det B = -460$. El producto de A por B es

$$AB = \begin{bmatrix} 17 & -13 \\ -21 & -11 \end{bmatrix}$$

El determinante de AB es $17(-11) - (-21)(-13) = -460$. Así, $\det(AB) = \det A \det B$.

b) El determinante de A es igual a

$$1\begin{vmatrix} 1 & 8 \\ 7 & 6 \end{vmatrix} - 4\begin{vmatrix} 5 & 8 \\ 0 & 6 \end{vmatrix} - 3\begin{vmatrix} 5 & 1 \\ 0 & 7 \end{vmatrix}$$

$$= (1)(-50) - (4)(30) - (3)(35) = -275$$

El determinante de B es igual a

$$2\begin{vmatrix} 2 & 5 \\ 8 & 7 \end{vmatrix} - 3\begin{vmatrix} -1 & 5 \\ 4 & 7 \end{vmatrix} + 6\begin{vmatrix} -1 & 2 \\ 4 & 8 \end{vmatrix}$$

$$= (2)(-26) - (3)(-27) + (6)(-16) = -67$$

Luego, $\det A \det B = 18{,}425$. Por su parte, el producto de A por B es

$$\begin{bmatrix} -14 & -13 & 5 \\ 41 & 81 & 91 \\ 17 & 62 & 77 \end{bmatrix}$$

y así, el determinante de AB es

$$\begin{aligned} &-14\begin{vmatrix} 81 & 91 \\ 62 & 77 \end{vmatrix} + 13\begin{vmatrix} 41 & 91 \\ 17 & 77 \end{vmatrix} + 5\begin{vmatrix} 41 & 81 \\ 17 & 62 \end{vmatrix} \\ &= (-14)(595) + (13)(1610) + (5)(1165) \\ &= 18{,}425 \end{aligned}$$

Por tanto, $\det(AB) = \det A \det B$.

Ejemplo 47.
Consideremos, nuevamente, las matrices del ejemplo 46. Veamos que, en estos casos, $\det(A + B) \neq \det A + \det B$.

Solución

a) El determinante de A es igual a -10 y el determinante de B es igual a 46. Luego, $\det A + \det B = 36$. La suma de A y B es

$$A + B = \begin{bmatrix} 6 & 3 \\ -6 & 7 \end{bmatrix}$$

El determinante de $A+B$ es 60. Por tanto, $\det(A+B) \neq \det A + \det B$.

b) Como vimos en el ejemplo 46, el determinante de A es igual a -275 y el determinante de B es igual a -67. Luego, $\det A + \det B = -342$. La suma de A y B es

$$\begin{bmatrix} 3 & 7 & 3 \\ 4 & 3 & 13 \\ 4 & 15 & 13 \end{bmatrix}$$

El determinante de $A + B$ es igual a -324. Por tanto, $\det(A + B) \neq \det A + \det B$.

Nota 5.
Los determinantes (que tuvieron su origen en los “rollos de bambú” con que los chinos entre el 200 a. C. y el 100 a. C. resolvían sistemas de ecuaciones lineales) fueron desarrollados formalmente desde G. W. Leibniz (1693) hasta A. L. Cauchy (1813) pasando por G. Cramer (1750), J. L. Lagrange (1773), A. T. Vandermonde (1771) (a quien se le considera el fundador formal de la teoría de los determinantes), P. S. Laplace (1772) y K. F. Gauss (1801) (a quien, como dijimos, le debemos el acuñar la palabra “determinante”). Hoy en día se ven como *herramientas útiles, convenientes, y simplificadoras de notación*, en la solución de sistemas de ecuaciones lineales, además de otras características importantes que explicaremos más adelante. Curiosamente, la historia muestra que hasta bien entrado el siglo XX, el estudio de los determinantes, de lejos, adelantaba al estudio de las matrices, como puede corroborarse en una cantidad apreciable de textos de la época.

Teorema 13. ***(Estructura de las soluciones de un sistema lineal)***
Un sistema lineal $m \times n$, $AX = b$, tiene soluciones de la forma $X = z + w$ donde z es una solución al sistema lineal homogéneo $AX = 0$, y w es una solución particular del sistema original $AX = b$.

Demostración
En efecto, si z es una solución al sistema homogéneo, entonces $A(X-z) = AX - Az = b - 0 = b$; luego $w = X - z$ es todavía una solución de $AX = b$. ■

Ejemplo 48.
En el ejemplo 5 de la lección anterior mostramos que el sistema homogéneo de ecuaciones

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

nos lleva a la solución $(x, y, z) = t(1, -2, 1)$ donde $t \in \mathbb{R}$. Podemos observar que una solución particular del sistema

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

es

$$(x_0, y_0, z_0) = (-1, 1, 0)$$

Por lo tanto, la solución general del sistema es

$$(x, y, z) = \underbrace{(-1, 1, 0)}_{\text{Solución particular}} + \underbrace{t\ (1, -2, 1)}_{\text{Solución general de la homogénea}} \qquad t \in \mathbb{R}$$

Esta solución se puede obtener estudiando la matriz ampliada

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 1 \\ 7 & 8 & 9 & 1 \\ 10 & 11 & 12 & 1 \end{array}\right]$$

y aplicando, de nuevo, las operaciones fila que se asignaron al sistema homogéneo en el ejemplo 5 de la lección 1.

Ejercicios 6

1) Muestre que el valor del determinante

$$\begin{vmatrix} 29 & 26 & 22 \\ 25 & 31 & 27 \\ 63 & 54 & 46 \end{vmatrix}$$

es 132 utilizando las distintas propiedades de éste.

2) Calcule $\det A$ si

$$A = \begin{bmatrix} a_{11} & 4 & 8 & 8 \\ 0 & a_{22} & 2 & 2 \\ 0 & 0 & a_{33} & 6 \\ 0 & 0 & 0 & a_{44} \end{bmatrix}$$

¿Encuentra alguna característica particular en este determinante?

3) a) Si A es una matriz diagonal, pruebe que $\det A$ es el producto de los números sobre la diagonal.

   b) ¿Cuál será el determinante de una matriz triangular?

4) Si $\det A = \det B$, ¿será que $A = B$?

5) Si $A^2 = A$ (*matriz idempotente*), ¿será que $\det A = 0$ ó $\det A = 1$?

6) Si $A^T = -A$ (*matriz antisimétrica*), ¿será que $\det A = 0$?

7) ¿Cuál es el determinante de una matriz A que satisface $A^n = A$ para algún $n > 2$?

8) Muestre que si $A^T A = I_n$, entonces $\det A = 1$ ó $\det A = -1$.

9) Evalúe los determinantes de las siguientes matrices, y escríbalas como expresiones algebraicas simples:

a) $$\begin{bmatrix} -ab & ac & ae \\ bd & -cd & de \\ bf & cf & -ef \end{bmatrix}$$

b) $$\begin{bmatrix} a_1 + ka_2 + la_3 & a_2 + ma_3 & a_3 \\ b_1 + kb_2 + lb_3 & b_2 + mb_3 & b_3 \\ c_1 + kc_2 + lc_3 & c_2 + mc_3 & c_3 \end{bmatrix}$$

c) $$\begin{bmatrix} c & a & d & b \\ a & c & d & b \\ a & c & b & d \\ c & a & b & d \end{bmatrix}$$

d) $$\begin{bmatrix} 1 & p & q & r+s \\ 1 & q & r & p+s \\ 1 & r & s & p+q \\ 1 & s & p & q+r \end{bmatrix}$$

[Indicación: el determinante de b) es igual a $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$; y los determinantes c) y d) son iguales a 0]

10) Pruebe que

a) $\begin{vmatrix} 1 & a & a^2 - bc \\ 1 & b & b^2 - ac \\ 1 & c & c^2 - ab \end{vmatrix} = 0$ b) $\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ b+c & a+c & a+b \end{vmatrix} = 0$

Calcule también estos dos determinantes a la luz del resultado geométrico del área de un triángulo que estudiamos en el ejemplo 30. [Indicación: un segmento de recta no tiene área].

11) Encuentre la solución general del sistema $AX = b$ donde

a)

$$A = \begin{bmatrix} 8 & -10 \\ -4 & 5 \\ 12 & -15 \end{bmatrix}, \qquad b = \begin{bmatrix} 4 \\ -2 \\ 6 \end{bmatrix}$$

b)

$$A = \begin{bmatrix} 9 & 2 & 1 \\ -3 & 2 & 4 \\ 5 & 2 & 2 \\ 7 & -2 & -5 \end{bmatrix}, \qquad b = \begin{bmatrix} 23 \\ 5 \\ 17 \\ 1 \end{bmatrix}$$

En cada caso, exprese la solución general de la forma $X = z + w$ donde z es la solución general al sistema lineal homogéneo $AX = 0$, y w es una solución particular del sistema $AX = b$.

7. Contexto económico

a. Primer modelo lineal formal en la teoría económica: sobre las tasas de intercambio (Cournot (1838))

Augustin Louis Cournot [1801-1877] fue el pionero y eje central en la historia de los desarrollos matemáticos en economía. Su principal trabajo fue *Recherches sur les Principes Mathématiques de la Théorie des Richesses* (1838), al que le seguiría su versión literaria *Principes de la Théorie des Richesses* (1863), y en el último año de su vida aparecería *Revue Sommaire des Doctrines Économiques* (1877). Pero también escribiría sobre probabilidad y filosofía de la ciencia (*Exposition de la Théorie des Chances et des Probabilités* (1843); y *Matérialisme, Vitalisme, Rationalisme: Études des Données de la Science en Philosophie* (1875)).

Sin embargo, todos estos trabajos, y fundamentalmente el primero, tuvieron que esperar muchos años para recibir el reconocimiento que merecían; pero cuando aparecieron en publicaciones de economistas de la talla de William Jevons, Alfred Marshall, Francis Edgeworth y Leon Walras, movieron dramáticamente el curso de la teoría económica. Marshall, por ejemplo, reconoce a Cournot desde 1868 como un gran maestro y como fuente de inspiración en cuanto a forma de pensamiento; Jevons decía haber leído a Cournot en 1872 y se sorprendió de encontrar "un análisis maravilloso de las leyes de la oferta y la demanda, y de las relaciones de precios, producción, consumo, gasto y beneficios"; y Walras reconoce en Cournot a aquel que le allega "la idea de utilizar el cálculo de funciones en la elaboración de [su] doctrina". Sin embargo, Cournot tuvo que luchar por cuarenta años para que sus ideas fueran aceptadas, y lo hizo con persistencia y humor. Comprendió que se necesitaba otra generación para que se pudieran entender sus ideas. Después de su muerte, Jevons y Walras mostrarían que tenía razón.

En el capítulo III (*De los Intercambios*) de su *Researches* de 1838, Cournot presentaba el primer modelo lineal formal que conozca la historia del pensamiento económico. Allí destacaba la necesidad de medidas uniformes para el sistema monetario europeo "tan a menudo desconocida por el egoísmo y la mala fe de los gobiernos". Así describía el problema, que lo presentamos en forma completa sólo en virtud de su valor histórico. No sobra anotar aquí que para Cournot el *valor de cambio* era el *único fundamento de la riqueza de una economía*, y de allí su preocupación por las condiciones en que se dan los intercambios monetarios.

> Supongamos que todos los comerciantes han adoptado la misma unidad monetaria, por ejemplo, un gramo de plata, o, lo que es lo

mismo, que la razón (*ratio*) de cada unidad monetaria con respecto a un gramo de plata está permanentemente establecida. El conocimiento de estas razones forma una gran parte de lo que aquellos vinculados a los negocios llaman ciencia del intercambio. Esta ciencia, que podría resumirse en una tabla común, no debería distraer nuestra atención. En otras palabras, no nos interesan los intercambios nominales sino únicamente los reales; es decir, la razón entre los valores de intercambio basados en el peso del gramo de plata de acuerdo a como sea pagado en diferentes lugares. Es claro también que el costo de intercambio, o la diferencia, con respecto a la unidad, de la razón de intercambio, no puede exceder el costo de transporte de este peso en plata de un lugar a otro, cuando el libre comercio en metales preciosos se permite entre los dos sitios, o el costo de transporte más el costo de contrabando cuando su comercio lo prohíbe la ley. Para encontrar las *ecuaciones de intercambio*, supondremos, para comenzar, que el costo de intercambio es menor que el costo de transporte, o que el intercambio se da sin ningún transporte real de dinero, [y] sin ningún cambio en la distribución de los metales preciosos entre los dos centros comerciales.

Supongamos primero sólo dos centros de intercambio. Designemos por $m_{1,2}$ el total de las sumas que el centro 1 le debe anualmente al centro 2; y $m_{2,1}$ el total de las sumas que el centro 2 le debe anualmente al centro 1; por $c_{1,2}$ la tasa de cambio en el lugar 1 con respecto al lugar 2, o la cantidad de plata dada en el lugar 2 a cambio de un peso de plata expresada por 1 y pagable en el lugar 1.

Adoptando esta notación y partiendo de la hipótesis de que los dos lugares balancean sus cuentas sin costos de transporte en ninguna dirección, es claro que tendremos

$$m_{1,2}c_{1,2} = m_{2,1} \quad \text{o} \quad c_{1,2} = \frac{m_{2,1}}{m_{1,2}}$$

Tenemos en general

$$c_{2,1} = \frac{1}{c_{1,2}}$$

y en este caso particular

$$c_{2,1} = \frac{m_{1,2}}{m_{2,1}}$$

$(\cdots)$ Ahora supongamos cualquier número de lugares comunicados, y sea $m_{i,k}$ el total de las sumas que el lugar i le debe anualmente al lugar k; y $c_{i,k}$ el coeficiente de intercambio i a k. $(\cdots)$

Ahora es fácil encontrar tantas ecuaciones como centros de correspondencia haya, partiendo siempre de la hipótesis de que no hay transporte de dinero real entre centros, y que así lo que un centro le debe a los otros es, para este primer centro, precisamente el valor que todos los otros centros le deben a él.

De esta última consideración se obtienen las siguientes ecuaciones:

$$\begin{array}{rcl}
m_{2,1}c_{2,1} + m_{3,1}c_{3,1} + \cdots + m_{r,1}c_{r,1} &=& m_{1,2} + m_{1,3} + \cdots + m_{1,r} \\
m_{1,2}c_{1,2} + m_{3,2}c_{3,2} + \cdots + m_{r,2}c_{r,2} &=& m_{2,1} + m_{2,3} + \cdots + m_{2,r} \\
m_{1,3}c_{1,3} + m_{2,3}c_{2,3} + \cdots + m_{r,3}c_{r,3} &=& m_{3,1} + m_{3,2} + \cdots + m_{3,r} \\
\vdots \quad \vdots \qquad \vdots \quad \vdots \qquad \cdots \qquad \vdots \quad \vdots &=& \vdots \qquad \vdots \qquad \cdots \qquad \vdots \\
m_{1,r}c_{1,r} + m_{2,r}c_{2,r} + \cdots + m_{r-1,r}c_{r-1,r} &=& m_{r,1} + m_{r,2} + \cdots + m_{r,r-1}
\end{array}$$

$(\cdots)$ *y, de hecho, si colocamos*

$$\begin{array}{rcl}
c_{1,2} = \dfrac{1}{c_{2,1}}, \; c_{1,3} &=& \dfrac{1}{c_{3,1}}, \; \cdots \; c_{1,r} = \dfrac{1}{c_{r,1}} \\
c_{3,2} = c_{3,1} \times c_{1,2} &=& \dfrac{c_{3,1}}{c_{2,1}} \\
\vdots &=& \vdots \\
c_{r-1,r} &=& \dfrac{c_{r-1,1}}{c_{r,1}}
\end{array}$$

entonces las ecuaciones son

$$\begin{array}{rcl}
m_{2,1}c_{2,1} + m_{3,1}c_{3,1} + \cdots + m_{r,1}c_{r,1} &=& m_{1,2} + m_{1,3} + \cdots + m_{1,r} \\
m_{1,2} + m_{3,2}c_{3,1} + \cdots + m_{r,2}c_{r,1} &=& (\, m_{2,1} + m_{2,3} + \cdots + m_{2,r} \,)c_{2,1} \\
m_{1,3} + m_{2,3}c_{2,1} + \cdots + m_{r,3}c_{r,1} &=& (\, m_{3,1} + m_{3,2} + \cdots + m_{3,r} \,)c_{3,1} \\
\vdots \qquad \vdots \quad \vdots \qquad \cdots \qquad \vdots \quad \vdots &=& \vdots \qquad \vdots \qquad \cdots \qquad \vdots \\
m_{1,r} + m_{2,r}c_{2,1} + \cdots + m_{r-1,r}c_{r-1,1} &=& (\, m_{r,1} + m_{r,2} + \cdots + m_{r,r-1} \,)c_{r,1}
\end{array}$$

$(\cdots)$ así hay exactamente tantas ecuaciones diferentes como variables independientes.

Cuando sólo se consideran tres centros, las ecuaciones serían

$$\begin{array}{rcl}
m_{2,1}c_{2,1} + m_{3,1}c_{3,1} &=& m_{1,2} + m_{1,3} \\
m_{1,2} + m_{3,2}c_{3,1} &=& (\, m_{2,1} + m_{2,3} \,)c_{2,1} \\
m_{1,3} + m_{2,3}c_{2,1} &=& (\, m_{3,1} + m_{3,2} \,)c_{3,1}
\end{array}$$

De esto puede obtenerse que

$$c_{2,1} = \frac{m_{3,1}m_{1,2} + m_{1,2}m_{3,2} + m_{1,3}m_{3,2}}{m_{2,1}m_{3,1} + m_{2,1}m_{3,2} + m_{3,1}m_{2,3}}$$

$$c_{3,1} = \frac{m_{2,1}m_{1,3} + m_{1,2}m_{2,3} + m_{1,3}m_{2,3}}{m_{2,1}m_{3,1} + m_{2,1}m_{3,2} + m_{3,1}m_{2,3}}$$

y, por lo tanto,

$$c_{3,2} = \frac{m_{2,1}m_{1,3} + m_{1,2}m_{2,3} + m_{1,3}m_{2,3}}{m_{3,1}m_{1,2} + m_{1,2}m_{3,2} + m_{1,3}m_{3,2}}$$

La composición de los valores de $c_{2,1}$, $c_{3,1}$ y $c_{3,2}$ en este caso particular, muestra con suficiente claridad cómo la razón de $m_{1,2}$ a $m_{2,1}$ puede variar considerablemente sin causar grandes variaciones en el valor de $c_{2,1}$; o, en otras palabras, cómo las interconexiones de los centros de intercambio disminuyen las variaciones de la tasa de cambio de un lugar a otro (...).

Nota 6.
El lector podrá observar que Cournot no debía estar, en la década de 1830, familiarizado con el concepto de determinante, y la razón de esto es que por aquella época no se utilizaba ampliamente. Si lo hubiera tomado efectivamente, habría expresado la solución general del sistema de orden r, en lugar de restringirse al caso especial de orden 3.

Nota 7.
Cabe anotar aquí que ninguno de los grandes economistas hasta, incluso, principios del siglo XX (exceptuando a Walras, Jevons y Marshall) escribía con matemáticas y, por supuesto, tampoco intentaron describir los problemas económicos con la precisión de Cournot. Presentaban algunos casos numéricos que adornaban de teoría matemática utilizando, en ocasiones, formas funcionales específicas, pero era claro que no estaban tratando de construir ningún aparato matemático coherente. De otro lado, Cournot sí era alguien entrenado en matemáticas *cuyo conocimiento primario de la aplicación de éstas a otra disciplina era la aplicación a la Física donde la teoría está basada, en gran parte, en formas funcionales específicas.* Pero, hasta él, estas formas funcionales eran desconocidas en el estudio de las relaciones económicas: no se sabía, por ejemplo, cómo dar una forma específica a una función de demanda. Fue entonces muy imaginativo de Cournot el allegar la aproximación matemática *general* a la disciplina (hasta entonces literaria) de la economía.

Ejercicios complementarios

1) Encuentre el producto de

$$A = \begin{bmatrix} 1 & 4 & 1 \\ 0 & 1 & -2 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 3 & 1 & 1 \\ 7 & 0 & 1 \\ -1 & 1 & 2 \end{bmatrix}$$

2) Encuentre AB y BA si

$$A = \begin{bmatrix} 3 & -1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 7 \\ 2 \end{bmatrix}$$

3) Responda las siguientes preguntas:

   a) ¿Será que la suma de dos matrices diagonales del mismo tamaño es, de nuevo, una matriz diagonal?
   b) ¿Será que el producto de dos matrices diagonales compatibles es, de nuevo, una matriz diagonal?
   c) ¿Será que la suma de dos matrices triangulares superiores (inferiores) del mismo tamaño es, de nuevo, una matriz triangular superior (inferior)?
   d) ¿Será que el producto de dos matrices triangulares superiores (inferiores) compatibles es, de nuevo, una matriz triangular superior (inferior)?

4) Encuentre, si es posible, matrices 2×2 tales que

   a) $A^2 = -I_2$ b) $B^3 = 0$, pero $B \neq 0$

5) Sea

$$A = \begin{bmatrix} 1 & 4 \\ 1 & 5 \\ 1 & 6 \\ 1 & 7 \end{bmatrix}$$

   a) Calcule $P = A(\,A^T A\,)^{-1} A^T$ y $M = I_4 - P$.
   b) Verifique que $MP = 0$.

6) a) Pruebe que para toda matriz cuadrada A, la matriz $B = A + A^T$ es *simétrica*.
   b) Pruebe que la matriz $C = A - A^T$ es *antisimétrica*; es decir, $C^T = -C$.

c) Finalmente, pruebe que toda matriz A es la suma de una matriz simétrica B y una antisimétrica C.

7) Demuestre que si

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

conmuta con todas las matrices 2×2, entonces existe una constante k tal que $A = kI_2$.

8) Muestre que

a) $\begin{vmatrix} 67 & 19 & 21 \\ 39 & 13 & 14 \\ 81 & 24 & 26 \end{vmatrix} = -43$ b) $\begin{vmatrix} 29 & 26 & 22 \\ 25 & 31 & 27 \\ 63 & 54 & 46 \end{vmatrix} = 132$

9) Calcule el determinante (haciendo explícitos los cofactores) de las siguientes matrices:

a) $\begin{bmatrix} 4 & 1 \\ -2 & -5 \end{bmatrix}$ b) $\begin{bmatrix} 2 & -3 \\ 8 & 9 \end{bmatrix}$

c) $\begin{bmatrix} 7 & 6 & 4 \\ 3 & -2 & -1 \\ 3 & 5 & 6 \end{bmatrix}$ d) $\begin{bmatrix} -1 & 2 & 7 \\ -3 & 1 & 8 \\ -5 & 2 & 4 \end{bmatrix}$

e) $\begin{bmatrix} 6 & 7 & -3 & -4 \\ 5 & 5 & 1 & 0 \\ 2 & 0 & 4 & 3 \\ 1 & 2 & 2 & -5 \end{bmatrix}$ f) $\begin{bmatrix} 0 & 1 & 4 & 3 \\ 2 & 2 & 6 & -2 \\ -3 & 4 & 3 & 0 \\ 1 & 5 & 5 & 7 \end{bmatrix}$

10) Encuentre un ejemplo 4×4 en el cual

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} \neq (\det A)(\det D) - (\det B)(\det C)$$

donde A, B, C, D son matrices 2×2.

11) Si A, B y C son submatrices cuadradas 2×2, pruebe que

$$\det \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} = \det A \det B$$

¿Qué sucedería si A, B y C son submatrices 3×3? ¿Y $n \times n$?

12) ¿Será cierto que si una matriz cuadrada sólo tiene entradas ceros o unos, entonces su determinante es 1, 0 ó -1?

13) Encuentre el determinante de la matriz 2×2 definida por $A = [\,a_{ij}\,] = [i+j]$. Extienda al caso 3×3.

14) Calcule el área del triángulo cuyos vértices son $(1, 2)$, $(3, 4)$ y $(-1, 7)$. Dibuje una figura.

15) Calcule el área del paralelogramo generado por los puntos $(4, 1)$ y $(5, 2)$.

16) Pruebe que $(x_2 - x_1)(x_3 - x_1)(x_3 - x_2)$ es igual al determinante

$$\begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix}$$

A éste se le conoce como *determinante de (A. T.) Vandermonde* (1771) y surgió del problema de encontrar un polinomio cuadrático de la forma $y = a_2x^2 + a_1x + a_0$ tal que cruzara por los puntos (x_1, y_1), (x_2, y_2), (x_3, y_3), pues obsérvese que al sustituir estos puntos en la ecuación cuadrática obtenemos el sistema lineal

$$\begin{aligned} a_2x_1^2 + a_1x_1 + a_0 &= y_1 \\ a_2x_2^2 + a_1x_2 + a_0 &= y_2 \\ a_2x_3^2 + a_1x_3 + a_0 &= y_3 \end{aligned}$$

cuya solución (a_2, a_1, a_0) obliga la aparición de este determinante. ¿Bajo qué condiciones existe esta solución (a_2, a_1, a_0)?

17) a) Pruebe que

$$\begin{vmatrix} 1 & p & p^3 \\ 1 & q & q^3 \\ 1 & r & r^3 \end{vmatrix} = (p - q)(q - r)(r - p)(p + q + r)$$

b) Pruebe que

$$\begin{vmatrix} p^3 & p^2 & 1 \\ q^3 & q^2 & 1 \\ r^3 & r^2 & 1 \end{vmatrix} = -(q - r)(r - p)(p - q)(pq + qr + pr)$$

¿En qué forma generalizan los determinantes a) y b) anteriores, el problema original planteado por Vandermonde, según el ejercicio 16? ¿Cómo se relacionan estos determinantes con el determinante del área de un triángulo, discutido en el ejemplo 30 de esta lección?

18) Pruebe que

$$\begin{vmatrix} (b+c)^2 & ab & ac \\ ab & (c+a)^2 & bc \\ ac & bc & (a+b)^2 \end{vmatrix} = 2abc(a+b+c)^3$$

19) Pruebe que

a) $$\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix} \cdot \begin{vmatrix} b & -a & 0 \\ -a & 0 & b \\ 0 & b & -a \end{vmatrix} = \begin{vmatrix} 0 & bc-a^2 & b^2-ac \\ b^2-ac & 0 & bc-a^2 \\ bc-a^2 & b^2-ac & 0 \end{vmatrix}$$

b) $$\begin{vmatrix} l & m & n \\ m & n & l \\ n & l & m \end{vmatrix}^2 = \begin{vmatrix} l^2+m^2+n^2 & lm+mn+nl & ln+ml+nm \\ ml+nm+ln & m^2+n^2+l^2 & mn+nl+lm \\ nl+lm+mn & nm+ln+ml & n^2+m^2+l^2 \end{vmatrix}$$

20) Pruebe que

a) $$\begin{vmatrix} 6 & 8 & 7 & 11 \\ 13 & 9 & 12 & 4 \\ 7 & 10 & 3 & 5 \\ 14 & 1 & 6 & 15 \end{vmatrix} = 10{,}209$$

b) $$\begin{vmatrix} 2 & 1 & 9 & 1 \\ 4 & 3 & -1 & 2 \\ 1 & 4 & 3 & -2 \\ 3 & 2 & 1 & 4 \end{vmatrix} = 310$$

c) $$\begin{vmatrix} 1^2 & 2^2 & 3^2 \\ 2^2 & 3^2 & 4^2 \\ 3^2 & 4^2 & 5^2 \end{vmatrix} = -8$$

d) $$\begin{vmatrix} 1^2 & 2^2 & 3^2 & 4^2 \\ 2^2 & 3^2 & 4^2 & 5^2 \\ 3^2 & 4^2 & 5^2 & 6^2 \\ 4^2 & 5^2 & 6^2 & 7^2 \end{vmatrix} = 0$$

21) Pruebe que

$$\begin{vmatrix} 1+a_1 & 1 & 1 & 1 \\ 1 & 1+a_2 & 1 & 1 \\ 1 & 1 & 1+a_3 & 1 \\ 1 & 1 & 1 & 1+a_4 \end{vmatrix} = a_1\,a_2\,a_3\,a_4 \left[1+\frac{1}{a_1}+\frac{1}{a_2}+\frac{1}{a_3}+\frac{1}{a_4}\right]$$

22) ¿Qué propiedades *algebraicas* (teorema 2) tendría la multiplicación de matrices definida de la siguiente forma: si $A = (a_{ij})$ y $B = (b_{ij})$, entonces $AB \equiv (a_{ij} \cdot b_{ij})$? ¿Por qué cree el lector que ésta no ha sido elegida como multiplicación entre matrices?

23) El mismo problema anterior, pero ahora es la "división" de matrices: si $A = (a_{ij})$ y $B = (b_{ij})$, $b_{ij} \neq 0$ para todo i, j, ¿por qué cree el lector que no se ha definido la "división de matrices" utilizando la igualdad $\dfrac{A}{B} \equiv (\dfrac{a_{ij}}{b_{ij}})$?

* 24) Pruebe el teorema 7 para $n = 3$ y $n = 4$. [Indicación: tenga paciencia].

25) ¿Existirán matrices 3×1, B y C, tales que, simultáneamente,

$$B(1, 1, 0) = I_3$$
$$(1, 1, 0)C = [1]?$$

26) Encuentre el área del paralelogramo si los vértices en el plano xy son los siguientes:

a) $(-1, 2), (2, 0), (4, 3), (7, 1)$; b) $(2, -3), (5, -5), (4, -1), (1, 1)$

27) Encuentre el área del triángulo si los vértices en el espacio xyz son:

a) $(0, 0, 2), (0, 1, 2), (1, 1, 2)$; b) $(6, -1, 3), (6, 1, 1), (2, 2, 2)$

28) *Un problema de Cournot (1838).* En cierto sistema comercial hay tres centros de intercambio de bienes. Las mercancías fluyen de un centro a otro y cada centro utiliza una moneda diferente. Sean m_{ik} el total de las sumas anualmente debidas por el lugar i al lugar k, y c_{ik} el coeficiente monetario de intercambio de la moneda del lugar i a la moneda del lugar k. Entonces el sistema que determina las tasas de cambio es

$$\begin{aligned} m_{12} + m_{13} &= m_{21}c_{21} + m_{31}c_{31} \\ (m_{21} + m_{23})c_{21} &= m_{12} + m_{32}c_{31} \\ (m_{31} + m_{32})c_{31} &= m_{13} + m_{23}c_{21} \end{aligned}$$

donde las m_{ik}'s son conocidas y las c_{ij}'s son incógnitas. Escriba este sistema en forma matricial y, por eliminación gaussiana, pruebe que su solución es

$$c_{21} = \frac{m_{31}m_{12} + m_{12}m_{32} + m_{13}m_{32}}{m_{21}m_{31} + m_{21}m_{32} + m_{31}m_{23}}$$

$$c_{31} = \frac{m_{21}m_{13} + m_{12}m_{23} + m_{13}m_{23}}{m_{21}m_{31} + m_{21}m_{32} + m_{31}m_{23}}$$

¿Tendría el lector alguna interpretación económica de la solución anterior?

Lección 3

Sistemas de ecuaciones lineales: solución por matriz inversa

Introducción

La aplicación de un instrumento conciso como las matrices, trae simplicidad y lucidez a la solución de ecuaciones lineales pues abarca en fórmulas cortas y sintéticas (que nos recuerdan las del álgebra ordinaria) complicadas relaciones que conectan al conjunto de coeficientes del sistema. Para ilustrar esto, escribamos un sistema lineal $n \times n$ cualquiera:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad\quad \vdots \qquad \vdots \quad &\;\;\; \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned} \tag{1}$$

Si hacemos

$$A \equiv \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \; X \equiv \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ y } b \equiv \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

entonces el sistema lineal explícito (1) puede escribirse mediante la breve ecuación matricial

$$AX = b \tag{2}$$

Ahora: observemos en (2) que si existiera una única matriz C de tamaño $n \times n$ tal que $CA = I_n$, donde I_n es la matriz identidad de tamaño $n \times n$, entonces tendríamos que

$$C(\,AX\,) = Cb \quad \text{ó} \quad (\,CA\,)X = Cb \quad \text{ó} \quad I_nX = Cb$$

o

$$X = Cb$$

y así habríamos resuelto nuestro sistema (1) y la *solución sería única.* A esta matriz C la llamaremos, por razones obvias, la *matriz inversa* de la matriz A.

1. La matriz inversa

En lo que sigue exploramos la posibilidad de que, efectivamente, exista la matriz inversa y nos preguntaremos bajo qué condiciones es esto posible, además de las propiedades que esta matriz pueda tener.

Definición 1. (Matriz inversa (Cayley (1853)))
Sea A una matriz $n \times n$. Si existe una matriz C, también $n \times n$, tal que

$$AC = CA = I_n$$

entonces diremos que A es una *matriz invertible* (o *no-singular*) y que su *inversa* es C; esta matriz la denotamos por A^{-1}. Si tal matriz A^{-1} no existe, diremos que A es *no invertible*, o que es *singular*.

Nota 1 (La matriz inversa, si existe, es única).
Que podamos asignar esta notación unívoca (A^{-1}) se debe a que, de hecho, la matriz inversa, si existe, es única. En efecto, si C y D son inversas de A, entonces $CA = AD = I_n$; luego

$$C = CI_n = C(AD) = (CA)D = I_nD = D$$

Ejemplo 1.
Verifiquemos que:

a) Si $A = \begin{bmatrix} 1 & 1 \\ 7 & 8 \end{bmatrix}$ entonces $A^{-1} = \begin{bmatrix} 8 & -1 \\ -7 & 1 \end{bmatrix}$

b) Si $A = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}$ entonces $A^{-1} = \begin{bmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{bmatrix}$

c) Si $A = \begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & 4 \\ 0 & 0 & -5 \end{bmatrix}$ entonces $A^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{6} & -\frac{7}{30} \\ 0 & \frac{1}{3} & \frac{4}{15} \\ 0 & 0 & -\frac{1}{5} \end{bmatrix}$

d) Si $A = \begin{bmatrix} 1 & -3 & 0 & -2 \\ 3 & -12 & -2 & -6 \\ -2 & 10 & 2 & 5 \\ -1 & 6 & 1 & 3 \end{bmatrix}$ entonces $A^{-1} = \begin{bmatrix} 0 & 1 & 0 & 2 \\ 1 & -1 & -2 & 2 \\ 0 & 1 & 3 & -3 \\ -2 & 2 & 3 & -2 \end{bmatrix}$

Solución

Basta observar que

a) $AA^{-1} = \begin{bmatrix} 8-7 & -1+1 \\ 56-56 & -7+8 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = A^{-1}A$

b) $AA^{-1} = \begin{bmatrix} \frac{6}{7}+\frac{1}{7} & \frac{2}{7}-\frac{2}{7} \\ \frac{3}{7}-\frac{3}{7} & \frac{1}{7}+\frac{6}{7} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = A^{-1}A$

c) $$AA^{-1} = \begin{bmatrix} 1+0-0 & -\frac{1}{3}+\frac{1}{3}-0 & -\frac{7}{15}+\frac{4}{15}+\frac{1}{5} \\ 0+0+0 & 0+1+0 & 0+\frac{4}{5}-\frac{4}{5} \\ 0+0+0 & 0+0+0 & 0+0+1 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = A^{-1}A$$

d)
$$AA^{-1} = \begin{bmatrix} 0-3+0+4 & 1+3+0-4 & 0+6+0-6 & 2-6+0+4 \\ 0-12+0+12 & 3+12-2-12 & 0+24-6-18 & 6-24+6+12 \\ 0+10+0-10 & -2-10+2+10 & 0-20+6+15 & -4+20-6-10 \\ 0+6+0-6 & -1-6+1+6 & 0-12+3+9 & -2+12-3-6 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = A^{-1}A \quad \blacktriangle$$

Ahora mostramos las propiedades algebraicas de las matrices inversas que nos recuerdan (aunque no coinciden con) las propiedades algebraicas de los inversos de los números reales.

Teorema 1. ***(Álgebra de inversas)***
Sean A y B dos matrices $n \times n$ inversibles; entonces

a) I_n *es invertible e* $I_n^{-1} = I_n$

b) A^{-1} *también es invertible y además* $(A^{-1})^{-1} = A$

c) AB *es invertible y además* $(AB)^{-1} = B^{-1}A^{-1}$

d) $A + B$ *no es necesariamente invertible*

e) Si A *es invertible, entonces* A^T *también es invertible y* $(A^T)^{-1} = (A^{-1})^T$

f) *Si* $AC = AD$ *para ciertas matrices* C *y* D*, entonces* $C = D$

g) $det\ A^{-1} = \dfrac{1}{detA}$

Demostración

a) Puesto que $I_n I_n = I_n$ entonces $I_n^{-1} = I_n$.

b) Se deduce de la definición de matriz inversa $AA^{-1} = A^{-1}A = I_n$ y de la unicidad de la matriz inversa.

c) i) $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$

ii) $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n$

d) • Para ilustrar esto consideremos las siguientes matrices:

$$A = \begin{bmatrix} -1 & 2 \\ -6 & 14 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 1 \\ 10 & -2 \end{bmatrix}$$

Las matrices inversas de A y B son

$$A^{-1} = \begin{bmatrix} -7 & 1 \\ -3 & \frac{1}{2} \end{bmatrix} \qquad B^{-1} = \begin{bmatrix} \frac{1}{7} & \frac{1}{14} \\ \frac{5}{7} & -\frac{1}{7} \end{bmatrix}$$

Por tanto,

$$A^{-1} + B^{-1} = \begin{bmatrix} -\frac{48}{7} & \frac{15}{14} \\ -\frac{16}{7} & \frac{5}{14} \end{bmatrix}$$

Sin embargo, la inversa de

$$A + B = \begin{bmatrix} 1 & 3 \\ 4 & 12 \end{bmatrix}$$

no existe, porque si existiera $(A + B)^{-1} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

entonces de la igualdad $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 4 & 12 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ obtendríamos que

$$\begin{aligned} a_{11} + 4a_{12} &= 1 \\ a_{21} + 4a_{22} &= 0 \\ 3a_{11} + 12a_{12} &= 0 \\ 3a_{21} + 12a_{22} &= 1 \end{aligned}$$

y así, multiplicando la primera ecuación por 3, y comparándola con la tercera ecuación obtendríamos que

$$3a_{11} + 12a_{12} = 3$$
$$3a_{11} + 12a_{12} = 0$$

y estas no son compatibles.

- Veamos ahora un caso en el cual $(A + B)^{-1}$ sí existe, pero es diferente de $A^{-1} + B^{-1}$. Consideremos las siguientes matrices:

$$A = \begin{bmatrix} -1 & 2 \\ -5 & 14 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 1 \\ 10 & -2 \end{bmatrix}$$

Las matrices inversas de A y B son

$$A^{-1} = \begin{bmatrix} -\frac{7}{2} & \frac{1}{2} \\ -\frac{5}{4} & \frac{1}{4} \end{bmatrix} \qquad B^{-1} = \begin{bmatrix} \frac{1}{7} & \frac{1}{14} \\ \frac{5}{7} & -\frac{1}{7} \end{bmatrix}$$

Por tanto,

$$A^{-1} + B^{-1} = \begin{bmatrix} -\frac{47}{14} & \frac{4}{7} \\ -\frac{15}{28} & \frac{3}{28} \end{bmatrix}$$

Sin embargo,

$$(A + B)^{-1} = \begin{bmatrix} -4 & 1 \\ \frac{5}{3} & -\frac{1}{3} \end{bmatrix}$$

e) i) $A^T(A^{-1})^T = (A^{-1}A)^T = I_n^T = I_n$

ii) $(A^{-1})^T A^T = (AA^{-1})^T = I_n^T = I_n$

f) Si $AC = AD$, entonces $A^{-1}(AC) = A^{-1}(AD)$, y así $IC = ID$ ó $C = D$

g) Puesto que $AA^{-1} = I_n$ entonces, tomando determinantes, obtenemos que $(\det A)$ $(\det A^{-1}) = \det I_n$=1. De allí, el resultado se sigue inmediatamente. ■

Ejemplo 2.
Utilizando las matrices del ejemplo 1, resolvamos los siguientes sistemas:

a) $$\begin{aligned} x + y &= 4 \\ 7x + 8y &= 6 \end{aligned}$$

b) $$\begin{aligned} 2x - y &= 3 \\ x + 3y &= 5 \end{aligned}$$

c) $$\begin{aligned} 2x + y - z &= 2 \\ 3y + 4z &= 7 \\ -5z &= 6 \end{aligned}$$

d) $$\begin{aligned} x - 3y - 2w &= 5 \\ 3x - 12y - 2z - 6w &= 7 \\ -2x + 10y + 2z + 5w &= 10 \\ -x + 6y + z + 3w &= 8 \end{aligned}$$

Solución

a) Observemos que, en términos matriciales, el sistema es

$$\begin{bmatrix} 1 & 1 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

y, por tanto,

$$\begin{bmatrix} x \\ y \end{bmatrix} = A^{-1} \begin{bmatrix} 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 8 & -1 \\ -7 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 26 \\ -22 \end{bmatrix}$$

b) En términos matriciales, el sistema es

$$\begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$$

y, por tanto,

$$\begin{bmatrix} x \\ y \end{bmatrix} = A^{-1} \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{bmatrix} \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

c) El sistema, en forma matricial, es

$$\begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & 4 \\ 0 & 0 & -5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 7 \\ 6 \end{bmatrix}$$

y, por tanto,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1} \begin{bmatrix} 2 \\ 7 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{6} & -\frac{7}{30} \\ 0 & \frac{1}{3} & \frac{4}{15} \\ 0 & 0 & -\frac{1}{5} \end{bmatrix} \begin{bmatrix} 2 \\ 7 \\ 6 \end{bmatrix} = \begin{bmatrix} -\frac{47}{30} \\ \frac{59}{15} \\ -\frac{6}{5} \end{bmatrix}$$

d) Observemos que, en forma matricial, el sistema es

$$\begin{bmatrix} 1 & -3 & 0 & -2 \\ 3 & -12 & -2 & -6 \\ -2 & 10 & 2 & 5 \\ -1 & 6 & 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \\ 10 \\ 8 \end{bmatrix}$$

y, por consiguiente,

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = A^{-1} \begin{bmatrix} 5 \\ 7 \\ 10 \\ 8 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 2 \\ 1 & -1 & -2 & 2 \\ 0 & 1 & 3 & -3 \\ -2 & 2 & 3 & -2 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \\ 10 \\ 8 \end{bmatrix} = \begin{bmatrix} 23 \\ -6 \\ 13 \\ 18 \end{bmatrix}$$

Ejercicios 1

1) Pruebe que la matriz inversa de

a) $\begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix}$ es $\begin{bmatrix} 1 & 0 \\ -a & 1 \end{bmatrix}$; b) $\begin{bmatrix} 1 & 0 \\ 0 & a \end{bmatrix}$ es $\begin{bmatrix} 1 & 0 \\ 0 & a^{-1} \end{bmatrix}$ si $a \neq 0$;

c) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ es $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

2) Si $A = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}$, pruebe que $A^{-1} = \begin{bmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{bmatrix}$ y encuentre $(\, A^T \,)^{-1}$.

3) Si

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}$$

Utilice los resultados de 1) y 2) arriba para calcular $(\, AB \,)^{-1}$ y $(\, BA \,)^{-1}$.

4) Pruebe que la inversa de la matriz

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad \text{es} \quad A^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 1 & -1 \\ -3 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

5) Encuentre una matriz de tamaño 2×2, $A \neq 0$, tal que $A^2 = 0$. ¿Existirá una matriz invertible A tal que $A^2 = 0$?

6) Encuentre dos matrices A, B, de tamaño 2×2, tales que $(\, A + B \,)^2 \neq A^2 + 2AB + B^2$. ¿Qué condición sobre A y B se requiere para que la igualdad sí se tenga?

7) Dé ejemplos de matrices A *distintas* de la matriz cero y de la matriz idéntica, tales que $A^2 = A$.

8) a) Demuestre que si $A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$, entonces $A^2 = \begin{bmatrix} \cos 2\theta & -\operatorname{sen} 2\theta \\ \operatorname{sen} 2\theta & \cos 2\theta \end{bmatrix}$.

b) ¿Podría el lector calcular A^n para $n > 2$?

c) ¿Para qué valores de θ es A invertible? En tales casos, calcule A^{-1}.

[A esta matriz se le conoce como *"matriz de rotación"* por razones que entenderemos en la lección 6].

9) ¿Bajo qué condiciones sobre A y B se tendrá que $(A + B)^{-1} = A^{-1} + B^{-1}$?

10) Si A es invertible, ¿será A^n, para $n \geq 2$, invertible?; en tal caso, ¿cuál será $(A^n)^{-1}$?

11) Compare las propiedades algebraicas de la matriz inversa (teorema 1) con las propiedades algebraicas típicas de inversión de los números reales (Volumen 0 (Fundamentos)).

2. Cálculo de la matriz inversa mediante el método gaussiano

Que las matrices elementales subyacen al cálculo de matrices inversas se ve claro en el siguiente teorema:

Teorema 2. ***(Método gaussiano para el cálculo de la inversa)***

a) El producto de matrices elementales es una matriz elemental.

b) Toda matriz elemental es invertible y, además, esta inversa también es elemental.

c) Una matriz cuadrada es invertible si, y sólo si, puede reducirse a la matriz identidad mediante operaciones elementales.

d) Una matriz cuadrada es invertible si, y sólo si, es el producto de matrices elementales.

Demostración

a) Es consecuencia directa del teorema 3, lección 2.

b) i) Para $k \neq 0$, la inversa de la matriz

$$A = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & k & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$$

es

$$A^{-1} = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \frac{1}{k} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$$

ii) La inversa de la matriz

$$A = \begin{array}{cl} \begin{bmatrix} 1 & 0 & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \cdots & 0 \\ 0 & 1 & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \mathbf{0} & \cdots & \mathbf{1} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \mathbf{1} & \cdots & \mathbf{0} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \cdots & 1 \end{bmatrix} & \begin{array}{l} \\ \\ \\ \longleftarrow \text{Fila j} \\ \\ \longleftarrow \text{Fila i} \\ \\ \\ \end{array} \end{array}$$

es

$$A^{-1} = \begin{array}{cl} \begin{bmatrix} 1 & 0 & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \cdots & 0 \\ 0 & 1 & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \mathbf{0} & \cdots & \mathbf{1} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \mathbf{1} & \cdots & \mathbf{0} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \cdots & 1 \end{bmatrix} & \begin{array}{l} \\ \\ \\ \longleftarrow \text{Fila j} \\ \\ \longleftarrow \text{Fila i} \\ \\ \\ \end{array} \end{array}$$

iii) La inversa de la matriz

$$A = \begin{array}{cl} \begin{bmatrix} 1 & 0 & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \cdots & 0 \\ 0 & 1 & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \mathbf{1} & \cdots & \mathbf{0} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{k} & \cdots & \mathbf{1} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \cdots & 1 \end{bmatrix} & \begin{array}{l} \\ \\ \\ \longleftarrow \text{Fila j} \\ \\ \longleftarrow \text{Fila i} \\ \\ \\ \end{array} \end{array}$$

es

$$A^{-1} = \begin{array}{c} \begin{bmatrix} 1 & 0 & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \cdots & 0 \\ 0 & 1 & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \mathbf{1} & \cdots & \mathbf{0} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & -\boldsymbol{k} & \cdots & \mathbf{1} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \mathbf{0} & \cdots & \mathbf{0} & \cdots & 1 \end{bmatrix} \end{array} \begin{array}{l} \\ \\ \\ \longleftarrow \text{Fila j} \\ \\ \longleftarrow \text{Fila i} \\ \\ \\ \end{array}$$

c) Por el método gaussiano, A es invertible si, y sólo si, existen $E_1, \ldots, E_p$ tales que $(E_p \ldots E_1)A = I_n$.

d) De c), $A = (E_p \ldots E_1)^{-1} = E_1^{-1} \ldots E_p^{-1}$ ■

Ejemplo 3.
Este ejemplo ilustra el teorema anterior con respecto a que las matrices elementales son inversibles, que sus inversas son también matrices elementales, y que el producto de matrices elementales es también una matriz elemental. Para hacerlo, consideremos las siguientes matrices elementales:

$$A = \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix} \quad a \neq 0 \ , \qquad\qquad B = \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix}$$

Y observemos que

$$A^{-1} = \begin{bmatrix} \dfrac{1}{a} & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad B^{-1} = \begin{bmatrix} 1 & 0 \\ -a & 1 \end{bmatrix}$$

Además, notemos que

$$AB = \begin{bmatrix} a & 0 \\ a & 1 \end{bmatrix} \quad \text{y} \quad (AB)^{-1} = B^{-1}A^{-1} = \begin{bmatrix} \dfrac{1}{a} & 0 \\ -1 & 1 \end{bmatrix} \blacktriangle$$

Basados en el teorema 2, el algoritmo que expondremos a continuación nos debería permitir determinar si una matriz cuadrada es invertible y, también, calcular explícitamente la inversa de la matriz en el caso de que ésta exista. De hecho, este algoritmo no es más que una formalización del *método de eliminación gaussiana* ya estudiado en la lección 1 para encontrar las soluciones de un sistema de ecuaciones lineales. Veamos en qué consiste.

Dada una matriz cuadrada $A = [\,a_{ij}\,]_{n\times n}$, se podrían efectuar los siguientes pasos:

1. A partir de A construimos la matriz $A' = [\,A\,|\,I_{n\times n}\,]$; es decir,

$$A' = \left[\begin{array}{cccc|ccccc} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & 0 & \cdots & 1 \end{array}\right]$$

   Aquí a A' la llamaremos la *matriz ampliada* (o *aumentada*) de A.

2. Si $a_{11} = 0$, buscamos sobre la primera columna de A' una entrada no nula e intercambiamos la primera fila con la fila correspondiente del elemento no nulo encontrado. Si todos los elementos de la primera columna son nulos, entonces la matriz A no es invertible y el proceso termina. Si el elemento $a_{11} \neq 0$, dividimos toda la primera fila de A' por a_{11}.

3. Efectuamos operaciones elementales entre filas con el fin de hacer cero todas las posiciones de la primera columna que estén por debajo de la posición $(\,1, 1\,)$.

4. Si $a_{22} = 0$, buscamos sobre la segunda columna de A' un elemento distinto de cero que esté por debajo de la posición $(\,2, 2\,)$ e intercambiamos la segunda fila con la fila correspondiente del elemento no nulo encontrado. Si todos los elementos son nulos, la matriz A no es invertible y el proceso termina. Si el elemento $a_{22} \neq 0$ dividimos toda la segunda fila A' por a_{22}.

5. Efectuamos operaciones elementales entre filas con el fin de hacer cero todas las posiciones de la segunda columna que estén por *debajo* y por *encima* de la posición $(\,2, 2\,)$.

6. En general, si $a_{kk} = 0$, buscamos sobre la k-ésima columna de A' un elemento distinto de cero que esté por debajo de la posición $(\,k, k\,)$ e intercambiamos la k-ésima fila con la fila correspondiente al elemento no nulo encontrado. Si todos los elementos son nulos, la matriz A no es invertible y el proceso termina. Si el elemento $a_{kk} \neq 0$ dividimos toda la k-ésima fila de A' por a_{kk}.

7. Efectuamos operaciones elementales entre filas con el fin de hacer cero todas las posiciones de la k-ésima columna que estén por debajo y por encima de la posición $(\,k, k\,)$.

8. El proceso continúa así sucesivamente.

Si la matriz A posee inversa, ésta aparecerá en el lado derecho de la matriz aumentada, una vez haya aparecido la matriz identidad I_n en el lado izquierdo.

Ejemplo 4.
Calculemos la inversa, si existe, de la matriz

$$A = \begin{bmatrix} 0 & 3 & 12 \\ 3 & 9 & 6 \\ 1 & 3 & 9 \end{bmatrix}$$

Solución

1. Construimos la matriz aumentada

$$A' = \left[\begin{array}{ccc|ccc} 0 & 3 & 12 & 1 & 0 & 0 \\ 3 & 9 & 6 & 0 & 1 & 0 \\ 1 & 3 & 9 & 0 & 0 & 1 \end{array}\right] \begin{array}{l} F_1 \\ F_2 \\ F_3 \end{array}$$

donde hemos representado las filas 1, 2 y 3 de A' mediante la notación F_1, F_2 y F_3, respectivamente.

2. Como el elemento a_{11} de A' es nulo, debemos buscar sobre la primera columna una entrada no nula. Podemos elegir la entrada a_{21} ó a_{31}. Como $a_{31} = 1$, es más práctico elegir esta entrada. Ahora efectuamos el intercambio de filas $F_1 \leftrightarrow F_3$ para obtener la matriz aumentada

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 9 & 0 & 0 & 1 \\ 3 & 9 & 6 & 0 & 1 & 0 \\ 0 & 3 & 12 & 1 & 0 & 0 \end{array}\right] \begin{array}{l} F_1 \longleftrightarrow F_3 \\ \\ F_3 \longleftrightarrow F_1 \end{array}$$

3. Luego hacemos cero las posiciones por debajo de la posición $(1, 1)$ mediante la siguiente operación entre filas: $F_2 - 3F_1 \longleftrightarrow F_2$. Notemos que la posición $(3, 1)$ ya es cero. De esta manera obtenemos la siguiente matriz aumentada:

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 9 & 0 & 0 & 1 \\ 0 & 0 & -21 & 0 & 1 & -3 \\ 0 & 3 & 12 & 1 & 0 & 0 \end{array}\right] \quad F_2 \longleftrightarrow F_2 - 3F_1$$

4. Como $a_{22} = 0$, buscamos por debajo de este elemento un elemento no nulo. En este caso $a_{32} = 3$. Efectuamos el intercambio de filas $F_2 \leftrightarrow F_3$, con el fin de obtener

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 9 & 0 & 0 & 1 \\ 0 & 3 & 12 & 1 & 0 & 0 \\ 0 & 0 & -21 & 0 & 1 & -3 \end{array}\right] \begin{array}{l} \\ F_2 \longleftrightarrow F_3 \\ F_3 \longleftrightarrow F_2 \end{array}$$

Ahora el elemento $a_{22} \neq 0$. Colocamos un 1 en la posición $(2, 2)$, efectuando la operación $\frac{1}{3}F_2$ obtenemos

$$\left[\begin{array}{rrr|rrr} 1 & 3 & 9 & 0 & 0 & 1 \\ 0 & 1 & 4 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & -21 & 0 & 1 & -3 \end{array}\right] \quad F_2 \longleftrightarrow \frac{1}{3}\,F_2$$

5. Hacemos cero las posiciones que están por debajo y por encima de la posición $(2, 2)$. En nuestro caso, basta con hacer cero la posición $(1, 2)$, efectuando la operación $F_1 - 3F_2 \longleftrightarrow F_1$, de lo cual obtenemos

$$\left[\begin{array}{rrr|rrr} 1 & 0 & -3 & -1 & 0 & 1 \\ 0 & 1 & 4 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & -21 & 0 & 1 & -3 \end{array}\right] \quad F_1 \longleftrightarrow F_1 - 3F_2$$

6. Como $a_{33} \neq 0$, podemos efectuar la operación $-\frac{1}{21}F_3$ con el fin de obtener un 1 en la posición $(3, 3)$. Así,

$$\left[\begin{array}{rrr|rrr} 1 & 0 & -3 & -1 & 0 & 1 \\ 0 & 1 & 4 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 1 & 0 & -\frac{1}{21} & \frac{1}{7} \end{array}\right] \quad F_3 \longleftrightarrow -\frac{1}{21}\,F_3$$

7. Hacemos cero las posiciones que están por encima de la posición $(3, 3)$, efectuando las siguientes operaciones: $F_1 + 3F_3$ y $F_2 - 4F_3$. De esta manera obtenemos la matriz aumentada

$$\left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -1 & -\frac{1}{7} & \frac{10}{7} \\ 0 & 1 & 0 & \frac{1}{3} & \frac{4}{21} & -\frac{4}{7} \\ 0 & 0 & 1 & 0 & -\frac{1}{21} & \frac{1}{7} \end{array}\right] \quad \begin{array}{l} F_1 \longleftrightarrow F_1 + 3F_3 \\ F_2 \longleftrightarrow F_2 - 4F_3 \end{array}$$

El anterior proceso nos indica que

$$A^{-1} = \begin{bmatrix} -1 & -\frac{1}{7} & \frac{10}{7} \\ \frac{1}{3} & \frac{4}{21} & -\frac{4}{7} \\ 0 & -\frac{1}{21} & \frac{1}{7} \end{bmatrix}$$

Para comprobar que los cálculos se hicieron correctamente, es recomendable verificar que $AA^{-1} = I_3$ y esto queda como ejercicio para el lector.

Ejemplo 5.
Encontremos la inversa de

$$A = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}$$

y resolvamos el sistema

$$\begin{aligned} -x + y + 2z &= 5 \\ 3x - y + z &= 7 \\ -x + 3y + 4z &= 8 \end{aligned}$$

Solución

$$[\,A\,|\,I\,] = \left[\begin{array}{rrr|rrr} -1 & 1 & 2 & 1 & 0 & 0 \\ 3 & -1 & 1 & 0 & 1 & 0 \\ -1 & 3 & 4 & 0 & 0 & 1 \end{array}\right] \begin{array}{l} F_1 \\ F_2 \\ F_3 \end{array}$$

$$\left[\begin{array}{rrr|rrr} -1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 7 & 3 & 1 & 0 \\ 0 & 2 & 2 & -1 & 0 & 1 \end{array}\right] \begin{array}{l} \\ F_2 \longleftrightarrow F_2 + 3F_1 \\ F_3 \longleftrightarrow F_3 - F_1 \end{array}$$

$$\left[\begin{array}{rrr|rrr} -1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 7 & 3 & 1 & 0 \\ 0 & 0 & -5 & -4 & -1 & 1 \end{array}\right] \begin{array}{l} \\ \\ F_3 \longleftrightarrow F_3 - F_2 \end{array}$$

$$\left[\begin{array}{rrr|rrr} 1 & -1 & -2 & -1 & 0 & 0 \\ 0 & 1 & 3.5 & 1.5 & 0.5 & 0 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array}\right] \begin{array}{l} F_1 \longleftrightarrow -F_1 \\ F_2 \longleftrightarrow \frac{1}{2}\,F_2 \\ F_3 \longleftrightarrow -\frac{1}{5}\,F_3 \end{array}$$

$$\left[\begin{array}{rrr|rrr} 1 & -1 & 0 & 0.6 & 0.4 & -0.4 \\ 0 & 1 & 0 & -1.3 & -0.2 & 0.7 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array}\right] \begin{array}{l} F_1 \longleftrightarrow F_1 + 2F_3 \\ F_2 \longleftrightarrow F_2 - 3.5F_3 \\ \end{array}$$

$$\left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -0.7 & 0.2 & 0.3 \\ 0 & 1 & 0 & -1.3 & -0.2 & 0.7 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array}\right] \begin{array}{l} F_1 \longleftrightarrow F_1 + F_2 \\ \\ \end{array}$$

Luego

$$A^{-1} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix}$$

que se puede comprobar realizando la multiplicación AA^{-1} y mostrando que ésta coincide con I_3. Por tanto, la solución al sistema

$$\begin{aligned} -x+y+2z&=5\\ 3x-y+z&=7\\ -x+3y+4z&=8 \end{aligned}$$

es

$$\begin{bmatrix} x\\ y\\ z \end{bmatrix} = A^{-1}\begin{bmatrix} 5\\ 7\\ 8 \end{bmatrix} = \begin{bmatrix} 0.3\\ -2.3\\ 3.8 \end{bmatrix}$$

Ejemplo 6.
Encontremos la inversa de

$$A = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3}\\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4}\\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}$$

Solución

$$[\,A\,|\,I\,] = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & | & 1 & 0 & 0\\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & | & 0 & 1 & 0\\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & | & 0 & 0 & 1 \end{bmatrix} \begin{array}{l} F_1\\ F_2\\ F_3 \end{array}$$

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & | & 1 & 0 & 0\\ 0 & \frac{1}{12} & \frac{1}{12} & | & -\frac{1}{2} & 1 & 0\\ 0 & \frac{1}{12} & \frac{4}{45} & | & -\frac{1}{3} & 0 & 1 \end{bmatrix} \begin{array}{l} F_2 \longleftrightarrow F_2 - \frac{1}{2}F_1\\ F_3 \longleftrightarrow F_3 - \frac{1}{3}F_1 \end{array}$$

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & | & 1 & 0 & 0\\ 0 & \frac{1}{12} & \frac{1}{12} & | & -\frac{1}{2} & 1 & 0\\ 0 & 0 & \frac{1}{180} & | & \frac{1}{6} & -1 & 1 \end{bmatrix} \quad F_3 \longleftrightarrow F_3 - F_2$$

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & | & 1 & 0 & 0\\ 0 & 1 & 1 & | & -6 & 12 & 0\\ 0 & 0 & 1 & | & 30 & -180 & 180 \end{bmatrix} \begin{array}{l} F_2 \longleftrightarrow 12F_2\\ F_3 \longleftrightarrow 180F_3 \end{array}$$

$$\begin{bmatrix} 1 & 0 & -\frac{1}{6} & | & 4 & -6 & 0\\ 0 & 1 & 1 & | & -6 & 12 & 0\\ 0 & 0 & 1 & | & 30 & -180 & 180 \end{bmatrix} \quad F_1 \longleftrightarrow F_1 - \frac{1}{2}F_2$$

$$\begin{bmatrix} 1 & 0 & 0 & | & 9 & -36 & 30\\ 0 & 1 & 0 & | & -36 & 192 & -180\\ 0 & 0 & 1 & | & 30 & -180 & 180 \end{bmatrix} \begin{array}{l} F_1 \longleftrightarrow F_1 + \frac{1}{6}F_3\\ F_2 \longleftrightarrow F_2 - F_3 \end{array}$$

Luego

$$A^{-1} = \begin{bmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{bmatrix}$$

que se puede comprobar realizando la multiplicación AA^{-1} y mostrando que ésta coincide con I_3.

Ejercicios 2

1) Calcule, utilizando el método gaussiano, la inversa (si existe) de las siguientes matrices:

a) $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & -2 \\ -1 & -2 & 3 \end{bmatrix}$ b) $\begin{bmatrix} 4 & 2 & -3 \\ 6 & 3 & 5 \\ 1 & 1 & 2 \end{bmatrix}$

c) $\begin{bmatrix} 2 & 2 & -1 \\ 1 & 1 & 1 \\ 2 & -4 & 3 \end{bmatrix}$ d) $\begin{bmatrix} 1 & 4 & 3 \\ 2 & 1 & 0 \\ 0 & 7 & 6 \end{bmatrix}$

2) Considere el siguiente sistema de ecuaciones lineales:

$$\begin{aligned} x + 2y + z &= q \\ x - y - z &= 1 \\ 2x + y + pz &= 0 \end{aligned}$$

donde p y q son parámetros. Responda las siguientes preguntas utilizando el método gaussiano:

a) ¿Para qué valores de p y q existe más de una solución?

b) ¿Para qué valores de p y q no existe solución?

c) ¿Para qué valores de p y q existe una única solución? ¿Cuál es esta solución?

3. Cálculo de la matriz inversa mediante determinantes (regla de Cramer)

Ya sabemos que el sistema (caso 2×2)

$$AX = b$$

donde

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \; X = \begin{bmatrix} x \\ y \end{bmatrix} \text{ y } b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

tiene solución única si, y sólo si,

$$\det A = a_{11}a_{22} - a_{21}a_{12} \neq 0$$

En este caso, es muy fácil probar que la matriz inversa A^{-1} existe:

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

y, por tanto, la solución está dada por

$$\begin{bmatrix} x \\ y \end{bmatrix} = A^{-1}b = \frac{1}{\det A} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \frac{1}{\det A} \begin{bmatrix} a_{22}b_1 - a_{12}b_2 \\ -a_{21}b_1 + a_{11}b_2 \end{bmatrix}$$

$$= \frac{1}{\det A} \left[\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}, \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix} \right]^T$$

o, equivalentemente,

$$x = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\det A}, \qquad y = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\det A}$$

que es la *regla de Cramer* para $n = 2$.

Ejemplo 7.

a) Si $A = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}$, entonces $\det A = 10$ y

$$A^{-1} = \frac{1}{10} \begin{bmatrix} 4 & -1 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 0.4 & -0.1 \\ -0.2 & 0.3 \end{bmatrix}$$

b) Si $A = \begin{bmatrix} -3 & 8 \\ -2 & 7 \end{bmatrix}$, entonces $\det A = -5$ y

$$A^{-1} = -\frac{1}{5} \begin{bmatrix} 7 & -8 \\ 2 & -3 \end{bmatrix} = \begin{bmatrix} -1.4 & 1.6 \\ -0.4 & 0.6 \end{bmatrix} \blacktriangle$$

De manera similar, el sistema (caso 3×3)

$$AX = b$$

donde

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \; X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \; \text{y} \;\; b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

tiene solución única si, y sólo si,

$$\begin{aligned} \det A &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} + a_{21}a_{32}a_{13} - a_{21}a_{12}a_{33} \\ &+ a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13} \neq 0 \end{aligned}$$

En este caso se puede probar (ver teorema 3 adelante) que la matriz inversa A^{-1} existe y que, además,

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix}$$

donde

$A_{11} =$ cofactor de a_{11} en $\det A$ (es decir, el determinante 2×2 (con su respectivo signo) que acompaña a a_{11} en la expansión de $\det A$ y que surge de eliminar la fila 1 y la columna 1).

$\vdots$

$A_{33} =$ cofactor de a_{33} en $\det A$ (es decir, el determinante 2×2 que acompaña a a_{33} en la expansión de $\det A$ y que surge de eliminar la fila 3 y la columna 3).

Más explícitamente,

$$A_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} \qquad A_{12} = -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} \qquad A_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$A_{21} = -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} \qquad A_{22} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} \qquad A_{23} = -\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$$

$$A_{31} = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \qquad A_{32} = -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \qquad A_{33} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

A la matriz *traspuesta* de la matriz de cofactores

$$\text{Adj}\, A = \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix}$$

se le llama la *matriz adjunta de A*. Por lo tanto, podemos escribir que

$$A^{-1} = \frac{1}{\det\, A} \text{Adj}\, A$$

De esta forma, la solución al sistema 3×3, $AX = b$, está dada por $X = \dfrac{1}{\det\, A}(\text{Adj}\, A)(b)$ que nos conduce (con algo de álgebra y paciencia) a:

$$x = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\det A}, \qquad y = \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\det A}, \qquad z = \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\det A}$$

que es la *regla de Cramer* para $n = 3$.

Nota 2. (**Sobre la regla de Cramer**)
Después de los chinos, fue Leibniz el primero en reconocer esta "armoniosa notación" de los determinantes al eliminar dos incógnitas de tres ecuaciones lineales, que ni Viete ni Descartes habían advertido. Ciertas formas de determinantes eran utilizadas ocasionalmente en la segunda mitad del siglo XVIII, pero fue casi cien años después del tiempo de Gabriel Cramer [1704–1752] (quien en su *Introduction a l'Analyse des Lignes Courbes Algebriques* de 1750 incluyera la regla general que lleva su nombre) que los matemáticos le darían importancia al papel que los determinantes jugaban dentro de la geometría analítica. Aun así, si ha de asignársele a alguien la responsabilidad por la adopción y difusión de la notación de determinantes en los libros de texto, ésta debe ser al matemático alemán Ludwig O. Hesse [1811–1874].

Ejemplo 8.
La matriz inversa de

$$A = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix} \qquad \text{es} \qquad A^{-1} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix}$$

pues $\det A = 10$ y, además,

$$A_{11} = \begin{vmatrix} -1 & 1 \\ 3 & 4 \end{vmatrix} = -7 \quad A_{12} = -\begin{vmatrix} 3 & 1 \\ -1 & 4 \end{vmatrix} = -13 \quad A_{13} = \begin{vmatrix} 3 & -1 \\ -1 & 3 \end{vmatrix} = 8$$

$$A_{21} = -\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 2 \quad A_{22} = \begin{vmatrix} -1 & 2 \\ -1 & 4 \end{vmatrix} = -2 \quad A_{23} = -\begin{vmatrix} -1 & 1 \\ -1 & 3 \end{vmatrix} = 2$$

$$A_{31} = \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix} = 3 \quad A_{32} = -\begin{vmatrix} -1 & 2 \\ 3 & 1 \end{vmatrix} = 7 \qquad A_{33} = \begin{vmatrix} -1 & 1 \\ 3 & -1 \end{vmatrix} = -2$$

Luego la solución al sistema

$$\begin{aligned} -x + y + 2z &= 2 \\ 3x - y + z &= 1 \\ -x + 3y + 4z &= 2 \end{aligned}$$

es

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -0.6 \\ -1.4 \\ 1.4 \end{bmatrix}$$

Obviamente, también hubiéramos podido hallar la solución de este sistema más directamente por la regla de Cramer, así:

$$x = \frac{\begin{vmatrix} 2 & 1 & 2 \\ 1 & -1 & 1 \\ 2 & 3 & 4 \end{vmatrix}}{\begin{vmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{vmatrix}} = \frac{-8+2+6+4-6-4}{10} = -0.6$$

$$y = \frac{\begin{vmatrix} -1 & 2 & 2 \\ 3 & 1 & 1 \\ -1 & 2 & 4 \end{vmatrix}}{\begin{vmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{vmatrix}} = \frac{-4-2+12+2+2-24}{10} = -1.4$$

$$z = \frac{\begin{vmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 2 \end{vmatrix}}{\begin{vmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{vmatrix}} = \frac{2-1+18-2+3-6}{10} = 1.4$$

Ejemplo 9.
Calculemos la matriz inversa de

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 5 \\ 0 & 1 & 3 \end{bmatrix}$$

y hallemos la solución del sistema

$$\begin{aligned} x + 2y + 3z &= 5 \\ 2y + 5z &= 1 \\ y + 3z &= 7 \end{aligned}$$

Solución

a) *Primer método.* El determinante de la matriz A es igual a 1; además,

$$A_{11} = \begin{vmatrix} 2 & 5 \\ 1 & 3 \end{vmatrix} = 1 \qquad A_{12} = -\begin{vmatrix} 0 & 5 \\ 0 & 3 \end{vmatrix} = 0 \qquad A_{13} = \begin{vmatrix} 0 & 2 \\ 0 & 1 \end{vmatrix} = 0$$

$$A_{21} = -\begin{vmatrix} 2 & 3 \\ 1 & 3 \end{vmatrix} = -3 \quad A_{22} = \begin{vmatrix} 1 & 3 \\ 0 & 3 \end{vmatrix} = 3 \qquad A_{23} = -\begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} = -1$$

$$A_{31} = \begin{vmatrix} 2 & 3 \\ 2 & 5 \end{vmatrix} = 4 \qquad A_{32} = -\begin{vmatrix} 1 & 3 \\ 0 & 5 \end{vmatrix} = -5 \quad A_{33} = \begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} = 2$$

Por tanto,

$$A^{-1} = \begin{bmatrix} 1 & -3 & 4 \\ 0 & 3 & -5 \\ 0 & -1 & 2 \end{bmatrix}$$

Luego la solución al sistema

$$\begin{aligned} x + 2y + 3z &= 5 \\ 2y + 5z &= 1 \\ y + 3z &= 7 \end{aligned}$$

es

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1} \begin{bmatrix} 5 \\ 1 \\ 7 \end{bmatrix} = \begin{bmatrix} 30 \\ -32 \\ 13 \end{bmatrix}$$

b) *Segundo método*. Es posible también hallar la solución del sistema directamente por la regla de Cramer, así:

$$x = \frac{\begin{vmatrix} 5 & 2 & 3 \\ 1 & 2 & 5 \\ 7 & 1 & 3 \end{vmatrix}}{1} = 30; \quad y = \frac{\begin{vmatrix} 1 & 5 & 3 \\ 0 & 1 & 5 \\ 0 & 7 & 3 \end{vmatrix}}{1} = -32; \quad z = \frac{\begin{vmatrix} 1 & 2 & 5 \\ 0 & 2 & 1 \\ 0 & 1 & 7 \end{vmatrix}}{1} = 13 \quad \blacktriangle$$

El caso general de la regla de Cramer es ahora inmediato de establecer:

Teorema 3. ***(Regla de Cramer (G. Cramer (1750)))***

i) Sea A una matriz $n \times n$ tal que $\det A \neq 0$. Entonces

$$A^{-1} = \frac{1}{\det A} Adj\ A$$

donde $Adj\ A$ es la "matriz adjunta" de A definida por

$$Adj\ A = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}$$

y A_{ij} es el determinante de la matriz que resulta al eliminar de A su fila i y su columna j, y multiplicar después por $(-1)^{i+j}$.

ii) Definamos las siguientes n matrices:

$$A_1 = \begin{bmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{bmatrix} \quad A_2 = \begin{bmatrix} a_{11} & b_1 & \cdots & a_{1n} \\ a_{21} & b_2 & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & b_n & \cdots & a_{nn} \end{bmatrix} \cdots$$

$$A_n = \begin{bmatrix} a_{11} & a_{12} & \cdots & b_1 \\ a_{21} & a_{22} & \cdots & b_2 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & b_n \end{bmatrix}$$

(es decir, A_i es la matriz obtenida de A sustituyendo la i-ésima columna por la matriz columna b). Entonces la única solución del sistema lineal $AX = b$ está dada por $X = (x_1, x_2, ..., x_n)$ donde, para $i = 1, 2, \ldots, n$,

$$x_i = \frac{\det A_i}{\det A}$$

Demostración

i) Primero, observemos (con un poco de cuidado) que por la misma definición de los cofactores A_{ij}, se tiene que para todo $i = 1, 2, 3, ..., n$; $j = 1, 2, 3, ..., n$, :

$$a_{i1}A_{i1}+a_{i2}A_{i2}+a_{i3}A_{i3}+...a_{in}A_{in}=\det A$$
$$\text{Si } i \neq j \text{ entonces } a_{i1}A_{j1}+a_{i2}A_{j2}+a_{i3}A_{j3}+...a_{in}A_{jn} = 0$$

pues, en este último caso, la matriz A tendría la fila i y la fila j iguales a $[a_{i1}, a_{i2}, ..., a_{in}]$, y así su determinante sería 0.

A partir de estos dos hechos, y de la definición de matriz adjunta, se tendrá entonces que

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} \det A & 0 & \cdots & 0 \\ 0 & \det A & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \det A \end{bmatrix} = \det A \; I_n$$

Y por tanto,

$$A^{-1} = \frac{\text{Adj } A}{\det A}$$

ii) Puesto que

$$X = A^{-1}b = \frac{1}{\det A}(\text{Adj } A)b$$

entonces

$$x_i = \frac{1}{\det A}[b_1A_{1i} + b_2A_{2i} + ... + b_nA_{ni}] = \frac{1}{\det A} \det A_i \quad \blacksquare$$

¿Y qué hemos aprendido hasta ahora con respecto a las soluciones de un sistema de ecuaciones lineales $AX = b$?

Teorema 4. ***(¿Cuándo un sistema lineal tiene solución única?)***
Sea A una matriz $n \times n$. Entonces las siguientes tres afirmaciones son equivalentes:

a) *El sistema lineal* $AX = b$ *tiene solución única.*

b) A *es invertible.*

c) $\det A \neq 0$.

[En particular, si $b = 0$*, la solución del sistema homogéneo es única:* $X = 0$*].*

Demostración

Es inmediata a partir de la regla de Cramer (teorema 3). ■

a. Determinantes de matrices particionadas

El determinante de una matriz particionada diagonal en bloques cuadrados del mismo tamaño se obtiene de manera análoga al determinante de una matriz diagonal:

$$\begin{vmatrix} A_{11} & 0 \\ 0 & A_{22} \end{vmatrix} = |\,A_{11}\,| \cdot |\,A_{22}\,|$$

En el caso de una matriz particionada general en bloques cuadrados del mismo tamaño, su determinante es igual a

$$\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} = |\,A_{11}\,| \cdot |\,A_{22} - A_{21}A_{11}^{-1}A_{12}\,| = |\,A_{22}\,| \cdot |\,A_{11} - A_{12}A_{22}^{-1}A_{21}\,|$$

siempre y cuando la matriz A_{11} (en el primer caso) o la matriz A_{22} (en el segundo caso) sean invertibles. La prueba de esto queda como ejercicio al lector. Ilustrarlo con una matriz 4×4 y cuatro submatrices 2×2 podría ser suficiente para que el lector se convenza, en este nivel, de la veracidad de esta afirmación.

Ejemplo 10.
Encontremos el determinante de la matriz particionada

$$A = \begin{bmatrix} \begin{bmatrix} 2 & 4 \\ 0 & -5 \end{bmatrix} & \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \\ \begin{bmatrix} 6 & 3 \\ -2 & -3 \end{bmatrix} & \begin{bmatrix} 5 & 1 \\ 1 & 2 \end{bmatrix} \end{bmatrix}$$

donde

$$A_{11} = \begin{bmatrix} 2 & 4 \\ 0 & -5 \end{bmatrix} \quad A_{12} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \quad A_{21} = \begin{bmatrix} 6 & 3 \\ -2 & -3 \end{bmatrix} \quad A_{22} = \begin{bmatrix} 5 & 1 \\ 1 & 2 \end{bmatrix}$$

Solución

Tenemos que

$$|A_{22}| = 10 - 1 = 9, \qquad A_{22}^{-1} = \begin{bmatrix} \frac{2}{9} & -\frac{1}{9} \\ -\frac{1}{9} & \frac{5}{9} \end{bmatrix}$$

de forma tal que

$$A_{12}A_{22}^{-1}A_{21} = \begin{bmatrix} -\frac{16}{9} & -\frac{18}{9} \\ -2 & -3 \end{bmatrix}$$

$$A_{11} - A_{12}A_{22}^{-1}A_{21} = \begin{bmatrix} \frac{34}{9} & 6 \\ 2 & -2 \end{bmatrix}$$

$$\text{y} \quad |A_{11} - A_{12}A_{22}^{-1}A_{21}| = \frac{34}{9}(-2) - 6(2) = -\frac{176}{9}$$

Así, el determinante de la matriz particionada es $|A| = 9\left(-\dfrac{176}{9}\right) = -176.$

b. Inversas de matrices particionadas

La inversa de una matriz particionada diagonal en bloques cuadrados del mismo tamaño se obtiene de la siguiente manera:

$$\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & A_{22}^{-1} \end{bmatrix}$$

Y en el caso de una matriz particionada general en bloques cuadrados del mismo tamaño, una forma de matriz particionada inversa es

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1}(\,I + A_{12}F_2A_{21}A_{11}^{-1}\,) & -A_{11}^{-1}A_{12}F_2 \\ -F_2A_{21}A_{11}^{-1} & F_2 \end{bmatrix}$$

donde

$$F_2 = (\,A_{22} - A_{21}A_{11}^{-1}A_{12}\,)^{-1}$$

y la prueba de esto es inmediata al multiplicar a ambos lados de la matriz $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ por la matriz inversa indicada arriba.

Ejemplo 11.
Encontremos la inversa de la matriz particionada

$$A = \begin{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} & \begin{bmatrix} 3 & 4 \\ 2 & 2 \end{bmatrix} \\ \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} & \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix} \end{bmatrix}$$

Solución

Con la notación de arriba tenemos que

$$F_2 = \begin{bmatrix} -\frac{1}{20} & -\frac{3}{10} \\ -\frac{3}{20} & \frac{1}{10} \end{bmatrix}$$

$$A_{11}^{-1}(I + A_{12}F_2A_{21}A_{11}^{-1}) = \begin{bmatrix} -\frac{1}{4} & -\frac{3}{4} \\ -\frac{2}{5} & \frac{3}{10} \end{bmatrix}$$

$$-A_{11}^{-1}A_{12}F_2 = \begin{bmatrix} \frac{3}{4} & \frac{1}{2} \\ \frac{1}{5} & \frac{1}{5} \end{bmatrix} \quad ; \quad -F_2A_{21}A_{11}^{-1} = \begin{bmatrix} \frac{7}{20} & \frac{1}{20} \\ \frac{1}{20} & \frac{3}{20} \end{bmatrix}$$

Luego,

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} -\frac{1}{4} & -\frac{3}{4} & \frac{3}{4} & \frac{1}{2} \\ -\frac{2}{5} & \frac{3}{10} & \frac{1}{5} & \frac{1}{5} \\ \frac{7}{20} & \frac{1}{20} & -\frac{1}{20} & -\frac{3}{10} \\ \frac{1}{20} & \frac{3}{20} & -\frac{3}{20} & \frac{1}{10} \end{bmatrix}$$

Ejemplo 12.
Encontremos la inversa de la matriz particionada

$$A = \begin{bmatrix} \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix} & \begin{bmatrix} 2 & 0 \\ 5 & 1 \end{bmatrix} \\ \begin{bmatrix} 0 & 3 \\ 5 & 4 \end{bmatrix} & \begin{bmatrix} 2 & 4 \\ 1 & 1 \end{bmatrix} \end{bmatrix}$$

Solución

Aquí, tenemos que

$$F_2 = \begin{bmatrix} \frac{1}{8} & -\frac{7}{16} \\ \frac{9}{8} & -\frac{31}{16} \end{bmatrix}$$

$$A_{11}^{-1}(I + A_{12}F_2A_{21}A_{11}^{-1}) = \begin{bmatrix} \frac{7}{2} & -\frac{11}{8} \\ -\frac{11}{2} & \frac{17}{8} \end{bmatrix}$$

$$-A_{11}^{-1}A_{12}F_2 = \begin{bmatrix} \frac{3}{4} & -\frac{13}{8} \\ -\frac{5}{4} & \frac{23}{8} \end{bmatrix}; \quad -F_2A_{21}A_{11}^{-1} = \begin{bmatrix} \frac{3}{4} & -\frac{1}{16} \\ \frac{15}{4} & -\frac{25}{16} \end{bmatrix}$$

Por lo tanto,

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{7}{2} & -\frac{11}{8} & \frac{3}{4} & -\frac{13}{8} \\ -\frac{11}{2} & \frac{17}{8} & -\frac{5}{4} & \frac{23}{8} \\ \frac{3}{4} & -\frac{1}{16} & \frac{1}{8} & -\frac{7}{16} \\ \frac{15}{4} & -\frac{25}{16} & \frac{9}{8} & -\frac{31}{16} \end{bmatrix}$$

Ejercicios 3

1) En los siguientes casos calcule, si existe, la inversa de A y, a partir de ella, encuentre la solución al sistema lineal $AX = b$, donde $b = \begin{bmatrix} 7 \\ -1 \end{bmatrix}$, $X = \begin{bmatrix} x \\ y \end{bmatrix}$:

   a) $A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$ b) $A = \begin{bmatrix} 6 & -3 \\ 2 & -1 \end{bmatrix}$

   c) $A = \begin{bmatrix} 4 & 2 \\ -3 & -8 \end{bmatrix}$ d) $A = \begin{bmatrix} 12 & 9 \\ 15 & -1 \end{bmatrix}$

   Corrobore el resultado mediante la regla de Cramer.

2) Calcule, si existe, la inversa de las siguientes matrices:

   a) $A = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}$ b) $A = \begin{bmatrix} 7 & 9 & 11 \\ 8 & -8 & 5 \\ 4 & 60 & 30 \end{bmatrix}$

   c) $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ d) $A = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 1 & 8 \\ 0 & 4 & 9 \end{bmatrix}$

   y resuelva el sistema $AX = b$, donde $b = \begin{bmatrix} 0 \\ 1 \\ -3 \end{bmatrix}$, $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$. Corrobore el resultado mediante la regla de Cramer.

3) Muestre que el sistema de ecuaciones lineales

$$\begin{aligned} x_1 - x_4 &= 7 \\ 2x_2 + x_3 &= 2 \\ 4x_1 - x_2 &= -3 \\ 3x_3 - 5x_4 &= 2 \end{aligned}$$

tiene una única solución y halle esto utilizando el método que considere más conveniente.

4) Encuentre, si existe, A^{-1} para

$$A = \begin{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} & \begin{bmatrix} 3 & 2 \\ 4 & 1 \end{bmatrix} \\ \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix} & \begin{bmatrix} -1 & 7 \\ 1 & -1 \end{bmatrix} \end{bmatrix}$$

4. Contexto económico

a. Una "visión lineal" en la teoría del valor: la teoría de la imputación de von Wieser (1889)

La *teoría del valor*, o el estudio del *valor intrínseco* de una mercancía, tiene una muy larga tradición desde (por lo menos) Aristóteles quien lo estudió a partir de los conceptos de *valor de uso* y *valor de intercambio*. El valor de uso es la capacidad de una mercancía para satisfacer las necesidades humanas; y el valor de intercambio es el valor de la mercancía en términos de su capacidad de ser intercambiada por otra mercancía.

Para la *economía clásica* la existencia de *valor de uso* era requisito para tener valor de intercambio; es decir, una mercancía debía ser *útil* para que pudiera ser intercambiada. Y el valor de intercambio lo determinaban los costos de producir esa mercancía (salarios, beneficios y rentas). Por lo tanto, algunos de los economistas clásicos como Ricardo se preguntaban sobre la posible existencia de una *medida invariante* de valor.

Para los economistas de finales del siglo XIX (C. Menger, F. von Wieser y E. Böhm-Bawerk (Escuela austriaca); L. Walras, V. Pareto (Escuela de Lausanne, Suiza); W. Jevons, F. Edgeworth, A. Marshall (en Inglaterra)), aunque con algunas diferencias esenciales, el esfuerzo se condujo, ya no hacia el valor de uso sino hacia el *valor de intercambio a través de los precios de mercado* y así la teoría del valor devino en una teoría de la distribución de recursos escasos para usos específicos.

Friedrich von Wieser [1851-1926] fue un importante miembro de la Escuela Austriaca junto con Menger y Böhm-Bawerk (y que posteriormente continuaría con L. von Mises, F. A. Hayek y J. A. Schumpeter). Sus dos más importantes contribuciones fueron el desarrollo de la *teoría de la imputación*, en donde mostraba que los precios de los factores eran determinados por los precios de los productos (y no al contrario, como aseguraban los clásicos); y el de la *teoría del "costo de oportunidad"* (el costo asociado a la utilización de los factores en su mejor uso alternativo), en la que fundamentaba la teoría del valor, y que afirma que los precios relativos reflejan oportunidades perdidas. Puede decirse que *von Wieser apuntaló la teoría económica en el estudio de la distribución de recursos escasos basado en estos dos principios.*

A von Wieser se le menciona por dos trabajos, principalmente: *Natural Value* (1889) en el que describe en detalle la doctrina de los costos de oportunidad y la teoría de la imputación; y su *Social Economics* (1918) que es el intento ambicioso por aplicar aquella teoría al mundo real.

En el capítulo *V* de su *Natural Value ("The Principle of Solution. The Productive Contribution")*, von Wieser, en contraposición a su maestro Carl Menger, plantea una aproximación alternativa al proceso de valoración que dio en llamar *"la contribución productiva"*. Para Menger el valor de un bien dentro de un proceso productivo podía determinarse *retirándolo* de la combinación que daba lugar al producto y haciendo la diferencia entre los dos valores. Pero von Wieser aseguraba que este procedimiento no era correcto y que podía surgir cierta *sobrevaloración*. Para remediar esto, sugirió que la "contribución productiva del factor" era la que debería conducir el proceso de valoración. Así, "el elemento decisivo no es la porción del rendimiento que se pierde a través de la pérdida de un bien, sino la que se asegura por su posesión" (*Natural Value p. 85*). Veamos entonces el texto original[1].

> Supongamos que la vida de un cazador depende de su último cartucho para matar a un tigre que lo acecha. Si falla, todo está perdido, y el rifle y el cartucho juntos tendrán un valor calculable exacto. El valor de juntos, el rifle y el cartucho, es igual al éxito del disparo, ni más ni menos. Tomados aisladamente, de otro lado, no hay manera de calcular su valor. Son dos cantidades desconocidas para las cuales sólo hay una ecuación. Vamos a llamarlas X y Y, y pongamos el resultado exitoso en 100; todo ello puede decirse como que el valor está en la ecuación $X + Y = 100$.
>
> Ahora, supongamos que un artista fuera a decorar una vasija de metal que causa gran admiración por su forma perfecta. Supongamos, además, que éste fuera el único artista que pudiera hacer un trabajo realmente artístico, y que éste fuera el único trabajo artístico conocido. Y supongamos que, además de la pieza de metal que había empleado, no podría tenerse ningún otro material de características similares, ni oro, ni plata, ni madera, ni barro, ni siquiera otro pedazo del mismo metal.
>
> Sería absolutamente imposible distinguir en el valor de la vasija entre el valor de la mano de obra y el valor del material. Las capacidades del artista que concibe y ejecuta, y lo apropiado del material que colocaba en sus manos y daba forma, serían considerados igualmente condiciones irreemplazables para el éxito. Si, bajo las condiciones económicas existentes en ese momento, averiguamos cómo se valora al artista y cómo se valora el material, se deberá a la influencia del intercambio [...]. Pues estos actos no son aislados, sino que se dan junto con muchos otros de la misma clase,

[1] Traducción del editor.

y pueden también compararse con ellos. El mismo metal, del cual el artista crea una vasija de gran valor artístico, sirve también para darle a artículos de uso ordinario un valor extraordinario. Concluimos de esto que el metal mismo sólo puede tener un valor limitado, y que sólo una pequeña porción del alto valor del producto artístico se debe a él, mientras que en mucha mayor medida se le debe al aporte del artista. Se confirmaría nuestra opinión si observáramos que todo trabajo del artista es altamente valorado. Pero si, al mismo tiempo, observamos que también trabaja con materiales tales como oro y piedras preciosas, y que éstas, por su parte, igualmente entregan un alto valor a todos los productos de los cuales forman parte, nos vemos forzados a la conclusión de que, a pesar de su talento, la mayor parte del valor de su producto no siempre pertenece al artista, y que, cuando emplea estos materiales, una parte altamente importante, si no la más importante, del valor debe ser adscrita a ellos. Ciertamente no deberíamos nunca considerar únicamente el poder artístico o el material por sí mismos, y tampoco deberíamos medir los efectos de lo que cada uno, independientemente, es capaz de producir. Todo factor productivo, si va a ser efectivo, debe combinarse con otros y unir su acción a la de otros; pero los elementos que están unidos a él pueden alterar el resultado, y este hecho hace posible distinguir el efecto específico de cada elemento aisladamente, justo como si, solo, estuviese activo.

Es posible no solo separar estos efectos aproximadamente, sino colocarlos en cifras exactas, tan pronto como recojamos y midamos todas las circunstancias importantes del problema; tales como la cantidad de los productos, su valor, y la cantidad de los medios de producción empleados. Si tomamos cuidadosamente estos hechos en cuenta, obtendremos cierto número de ecuaciones y estaremos en posición de hacer un cálculo confiable de cuánto agrega cada instrumento de producción aisladamente. Para colocar en la más corta y típica fórmula todo lo que hemos dicho, tenemos, por ejemplo, en lugar de la ecuación $X + Y = 100$, las siguientes:

$$\begin{aligned} X + Y &= 100 \\ 2X + 3Z &= 290 \\ 4Y + 5Z &= 590 \end{aligned}$$

Aquí $X = 40$, $Y = 60$ y $Z = 70$.

El número de ecuaciones individuales será el número de combinaciones productivas individuales que se lleven a cabo dentro del proceso de producción. En estas ecuaciones los factores combinados de producción, en un lado, y el valor adquirido conjuntamente en el otro lado, están puestos, unos contra el otro, como cantidades equivalentes. Si sumamos todas las ecuaciones, la cantidad total de producción en riqueza se mantendrá equivalente al valor total del rendimiento. Esta suma debe ser adscrita totalmente a los elementos productivos individuales. Para cada elemento se tiene una participación definida en el resultado. Y a esta parte no podría dársele un valor ni más alto ni más bajo, sin romper la equivalencia entre riqueza productiva y su rendimiento.

Es la parte del rendimiento, asignada al factor productivo individual, la cual usualmente se llama brevemente "rendimiento" del factor en cuestión. El rendimiento de la mano de obra, el rendimiento de la tierra, el rendimiento del capital, yo lo llamaré la "contribución productiva" [...], de manera que pueda siempre ser claro si estamos hablando del rendimiento como un todo, o de la participación del factor aislado en el rendimiento. La contribución productiva, entonces, es la parte del rendimiento a la que está confinado el trabajo del elemento productivo individual en el rendimiento total de la producción. La suma de todas las contribuciones productivas coincide exactamente con el valor del rendimiento total.

Apenas necesita decirse que, de hecho, raramente podemos hacer estos cálculos tan exactamente, y nunca tan comprehensivamente. Las ecuaciones sí están de hecho allí, y en cada caso el resultado productivo se estima de acuerdo con el estándar del más alto rendimiento alcanzable. Pero la construcción de las ecuaciones se hace frecuentemente apenas con un leve grado de exactitud; y la suma de todas las ecuaciones nunca se alcanza completamente, y así no puede hacerse la separación entre los elementos individuales. No menos estamos entonces tratando de afirmar que la adición y división siempre funcionen bien; sólo que, en lugar de calcular directamente, tratamos de obtener nuestro objetivo en forma un tanto circunstancial mediante un método de prueba. Los valores obtenidos en el caso individual se aplican, hasta donde sea posible, a otros casos, y se corrigen, el uno contra el otro, hasta que al final se alcanza la división correcta. Y esto se hace inconmensurablemente más fácil por el hecho de que ya poseemos, en valores productivos conocidos y autenticados, una clave para la división, y que ésta sólo requiere adaptarse a los cambios que emergen de tiempo en tiempo, y en ningún caso se buscará calcular de una sola vez la contribución productiva de la masa total de los bienes de producción; sólo se requiere calcular de nuevo las contribuciones de los miembros individuales, y para ellas una buena base se encuentra en los valores previamente encontrados. Nuevos cálculos requieren hacerse únicamente en aquellos factores de la producción donde los rendimientos alcanzables y sus valores crecen o caen. Esto da origen a nuevas ecuaciones para los factores en cuestión, bien con valores totales más favorables o menos favorables. Dependiendo de si es uno u otro, se extenderá o limitará la producción, y elementos productivos serán atraídos de otros sectores de la producción, o atraídos hacia ellas, hasta que se encuentre el plan de producción más favorable. La experiencia obtenida cuando se transfiere uno, y después otro elemento productivo, y observar el efecto de cada combinación sobre el valor del rendimiento, nos da suficiente información con respecto a la cantidad que los elementos individuales aportan al rendimiento total.

Ejercicios complementarios

1) Muestre que si $A = P^{-1}BP$ para alguna matriz P, entonces $\det A = \det B$.

2) Muestre que si A y B son inversibles, entonces $A^{-1}(A + B)B^{-1} = A^{-1}+B^{-1}$, y, por tanto, $A(A^{-1}+B^{-1})B = A+B$. Escriba una expresión para $(A + B)^{-1}$.

3) Encuentre la inversa de

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

4) Pruebe, utilizando el método gaussiano, que la matriz inversa de

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 3 \\ 1 & 3 & 4 \end{bmatrix}$$

es

$$A^{-1} = \begin{bmatrix} 3 & -5 & 3 \\ -1 & 3 & -2 \\ 0 & -1 & 1 \end{bmatrix}$$

5) Resuelva los siguientes sistemas:

a)
$$\begin{aligned} 2x - 3y &= 8 \\ 4x - 5y + w &= 15 \\ 3x + 8w &= 1 \end{aligned}$$

b)
$$\begin{aligned} x - y + z &= 0 \\ 2x + 3y - 5z &= 7 \\ 3x - 4y - 2z &= -1 \end{aligned}$$

mediante la matriz inversa y la regla de Cramer.

6) Pruebe que la inversa de una matriz diagonal (si existe) es también diagonal. ¿Cuáles son los números de la diagonal?

∗7) Pruebe que la inversa de una matriz triangular, si existe, es también triangular.

8) a) Pruebe que si

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

entonces

$$A^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}$$

b) Indique con precisión cuáles son las *matrices elementales* $E_1, E_2, \ldots, E_p$ tales que

$$E_1\, E_2\, \ldots\, E_p\, A = I_4$$

c) Confirme que $E_1\, E_2 \ldots E_p = A^{-1}$

9) ¿Para qué valores de a y b es no-invertible la matriz

$$A = \begin{bmatrix} 1 & 2 & 2 \\ a & 8 & 3 \\ 0 & b & 3 \end{bmatrix}?$$

10) ¿Para qué valores de a existe la inversa de

$$A = \begin{bmatrix} a & -2 \\ 3 & -7 \end{bmatrix}?$$

11) Pruebe que la inversa de la matriz

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix} \quad \text{es} \quad A^{-1} = \begin{bmatrix} \frac{1}{8} & \frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{5}{4} & -\frac{3}{4} & -\frac{1}{4} \end{bmatrix}$$

12) ¿Por qué la matriz

$$\begin{bmatrix} 2 & 1 & 4 & 6 \\ 0 & 3 & 8 & 5 \\ 0 & 0 & 0 & 7 \\ 0 & 0 & 0 & 9 \end{bmatrix}$$

no es invertible?

13) Pruebe que si A es simétrica e invertible, entonces A^{-1} también es simétrica.

14) Pruebe que $A^{-1} - (A+B)^{-1} = A^{-1}(A^{-1} + B^{-1})^{-1}A^{-1}$ cuando estas inversas existen.

15) Sabemos que si A es invertible, entonces

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det A \cdot \det(\, D - CA^{-1}B\,)$$

donde A, B, C y D son submatrices cuadradas. Utilizando este resultado calcule

$$\det \begin{bmatrix} 1 & 2 & | & 0 & 1 \\ 1 & 3 & | & 0 & 1 \\ - & - & - & - & - \\ 1 & 1 & | & 1 & 3 \\ 1 & 2 & | & 1 & 4 \end{bmatrix}$$

y pruebe que es 1.

16) Muestre que encontrar un polinomio de grado 2, es decir, de la forma $ax^2 + bx + c$, cuyo valor en $x = 1$ y en $x = 3$ es cero, y el valor en $x = 4$ es 6, se reduce a resolver el sistema de ecuaciones lineales

$$\begin{aligned} a + b + c &= 0 \\ 9a + 3b + c &= 0 \\ 16a + 4b + c &= 6 \end{aligned}$$

Encuentre entonces, si existen, los coeficientes a, b, c, por el método de la matriz inversa. (¿Recuerda los determinantes de Vandermonde explicados en los ejercicios complementarios de la lección 2?)

17) Encuentre el área del rectángulo que satisface las siguientes condiciones: Si se le quitan 2 cm a cada uno de sus lados, el área del nuevo rectángulo es $\frac{2}{5}$ veces la del original. Si aumenta en 10 cm uno de sus lados, el área del nuevo rectángulo es el doble del original; y si se le agregan 2 cm al otro lado, el área del nuevo rectángulo es $\frac{3}{2}$ veces la del original. Resuelva el problema lineal que aparece aquí, mediante el método de la matriz inversa.

18) a) (*Modelo lineal de producción*). Supongamos que tres industrias se interrelacionan de tal modo que sus producciones se utilizan a su vez como insumos, de acuerdo con la matriz

$$A = \begin{bmatrix} 0.45 & 0.35 & 0 \\ 0.6 & 0 & 0.25 \\ 0.2 & 0.12 & 0.9 \end{bmatrix} = [\, a_{jk}\,]$$

donde a_{jk} es la fracción de la producción de la industria k que es consumida (comprada) por la industria j. Sea p_j el precio cobrado por la industria j por su producción. El problema planteado es encontrar precios tales que, para cada industria, los gastos totales sean iguales a los ingresos totales. Muestre que esto es equivalente a encontrar $p = (p_1, p_2, p_3)^T$ tal que $Ap = p$ (con $p_1, p_2, p_3 \geq 0$) y resuelva, si existe solución.

b) Si las industrias *no* consumen la totalidad de la producción como en el modelo anterior, entonces en lugar de $Ap = p$ se tiene $x - Ax = y$, donde $x = (x_1, x_2, x_3)^T$ es lo producido, Ax es lo consumido por las industrias y $y = (y_1, y_2, y_3)^T$ es la producción excedente disponible para otros consumidores. Resuelva $x - Ax = y$ si $y = (0.1, 0.3, 0.1)^T$ y la matriz de consumo es

$$A = \begin{bmatrix} 0.2 & 0.5 & 0 \\ 0.2 & 0 & 0.4 \\ 0.2 & 0.5 & 0.6 \end{bmatrix}$$

c) Comente los resultados en a) y b).

Lección 4

Vectores

Introducción

Además de las razones desde el punto de vista analítico ya señaladas en las lecciones anteriores, también existieron necesidades desde la geometría que condujeron al desarrollo del álgebra lineal. En la primera mitad del siglo XVII surgió una rama de las matemáticas completamente nueva que ahora se conoce como *geometría analítica*(Volumen 0 (Fundamentos)), que establecía conexiones entre curvas en un plano y ecuaciones con dos incógnitas, y que quizás fue promovida por la transición en Europa a métodos más avanzados de producción que requerían del impulso de todas las ciencias y, en particular, de la mecánica.

Elipses y parábolas, cuyas propiedades geométricas como secciones cónicas ya eran muy bien conocidas por los antiguos griegos poco menos de 2.000 años antes, dejaron de ser únicamente parte del estudio de la geometría. Después de que Kepler (1619) descubriera que los planetas giran alrededor del sol en elipses, y que Galileo (1632) mostrara que una piedra lanzada al aire describía una parábola, fue necesario calcular *explícitamente* estas figuras. Todas estas preguntas dieron vida, además, a otra rama fundamental del análisis matemático: *el cálculo diferencial e integral* de Newton y Leibniz, incluyendo allí el estudio de ecuaciones diferenciales simples. Estos tres campos (geometría analítica, cálculo y ecuaciones diferenciales) cambiaron radicalmente a las matemáticas de aquel entonces.

A comienzos de la década de 1600, ya la idea de la geometría analítica estaba cerca. A dos de los más notables matemáticos de la época, Pierre de Fermat [1601–1665] y René Descartes [1596–1650] se les da el crédito como los “inventores” de la geometría analítica. Descartes quería crear un método que pudiera ser aplicado a la solución de *todos* los problemas de la geometría. Su teoría estaba basada en dos conceptos: el concepto de coordenadas y el concepto de representación gráfica (por medio de coordenadas) de cualquier ecuación con dos incógnitas dentro de un plano. La última parte de su *Discourse on the Method of Rightly Conducting the Reason and Seeking the True in the Sciences with Applications: Dioptrics, Meteorology and Geometry* de 1637, contiene una presentación completa, aunque un tanto confusa, de lo que hoy se conoce como *geometría analítica*.

Fermat, por su parte, sólo compuso un breve ensayo sobre la geometría analítica llamado *Ad Locos Planos et Solidos Isagoge* de 1629, que es un libro dedicado a la línea, al círculo y a las secciones cónicas. Fermat proponía allí remitir la teoría de los lugares geométricos a un análisis que fuera apropiado a tales problemas y que, según él, abriera el camino al estudio de lugares geométricos generales. Y fue Fermat quien introdujo la útil idea de *variable algebraica*. Su visión dio significado a ecuaciones en dos incógnitas (que antes habían sido rechazadas por la geometría) permitiendo a una de las variables tomar valores sucesivos a lo largo de una línea medida sobre un eje dado a partir de un punto inicial, y a la otra variable corresponder a los valores determinados por la primera variable. Fermat mostraba entonces cómo una ecuación podía definir cierta curva con respecto a un sistema coordenado dado.

Y aunque ni Fermat ni Descartes inventaron las coordenadas, ni fueron los primeros en utilizar representaciones gráficas, sus aportes constituyeron una contribución decisiva al desarrollo de las matemáticas que hoy conocemos.

1. El concepto de vector

Acuñado por William R. Hamilton en 1853, el concepto de *vector* (palabra que proviene del Latin *vehĕre* y que significa *transportar*) llega a la geometría analítica (y de allí al álgebra lineal) de la mano del matemático francés Joseph Louis Lagrange [1736–1813]. En su *Analytic Mechanics*, publicada en 1788, Lagrange "aritmetizó" fuerzas, velocidades y aceleraciones en la misma forma que Descartes y Fermat "aritmetizaron" los puntos. Esta idea de Lagrange posteriormente tomó la forma de la llamada *"teoría de vectores"* que ha probado ser de importante ayuda en física, mecánica y en la tecnología.

Aparte de la fuente física de la que surge la noción de vector, el uso de este concepto confiere a la geometría analítica (y en consecuencia a la teoría de las ecuaciones lineales) una gran simplicidad y claridad. Aquí, un vector en el plano se da mediante *dos* números que son sus proyecciones sobre los dos ejes coordenados; y cada pareja de números reales, a su vez, puede representarse geométricamente en forma de un vector en el plano. Similarmente para los vectores en el espacio como triplas de números, etc. Por su parte, en geometría y física se utilizan dos clases de cantidades: *escalares y vectores*. Un *escalar* es simplemente un número real cualquiera: longitud, temperatura y voltaje son cantidades escalares. Un *vector*, en su lugar, es un concepto que se determina no sólo por su magnitud sino por su dirección, y por ello se le llama también *flecha* o *segmento de recta dirigido*. Una fuerza y una velocidad son ejemplos de vectores.

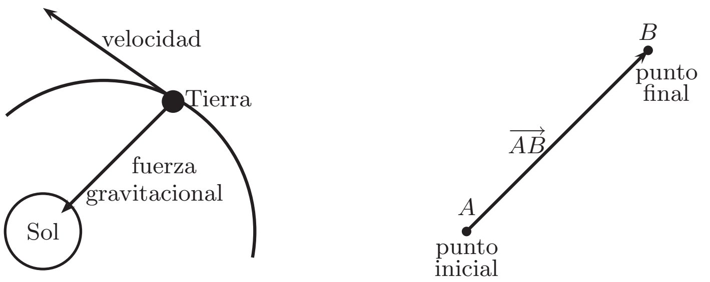


Figura 1

Desde esta última perspectiva, un vector (como flecha) tiene un *punto inicial* (cola de la flecha) y un *punto final* (cabeza de la flecha). A la longitud de un vector (la longitud de la flecha) se le llama la *norma del vector*. Dos vectores son iguales si tienen la misma longitud y la misma dirección. Desde esta perspectiva, *el punto inicial del vector puede escogerse de forma arbitraria.*

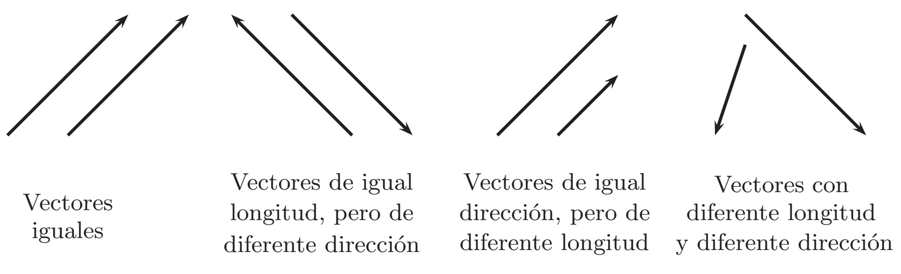


Figura 2

Ahora: si se elige un sistema de coordenadas XY en el plano, entonces el vector con punto inicial $A(\,x_1, y_1\,)$ y punto final $B(\,x_2, y_2\,)$ se denota, normalmente, $\overrightarrow{AB}$ (figura 3).

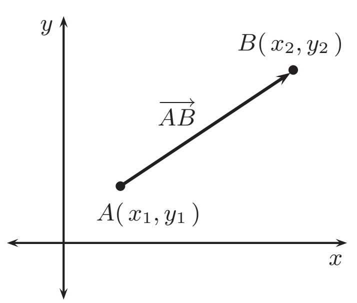


Figura 3

Pero como el punto inicial del vector puede escogerse arbitrariamente, podemos trasladarlo al origen $(0,0)$ y tener ahora el vector $A(a_1, a_2)$ (que se puede notar A ó $\overrightarrow{A}$) donde $a_1 \equiv x_2 - x_1,\ a_2 \equiv y_2 - y_1$ (figura 4).
Se observa entonces que partiendo de la definición geométrica (y física) de vector (flecha), se llega a su caracterización algebraica: vectores con punto inicial $(0,0)$. Esto ha permitido generalizar la idea de vector a más de dos dimensiones de la siguiente manera:

Definición 1. (Vector (Lagrange (1788), Grassmann (1844)))
Para cualquier $n = 1, 2, 3, \ldots$, a los elementos del conjunto

$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n)/x_i \in \mathbb{R}, i = 1, 2, \ldots, n\}$$

se les llamará *vectores* (de n dimensiones).

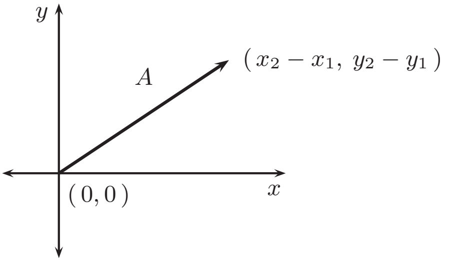


Figura 4

Nota 1.
Conviene destacar que, con esta definición, los vectores de n dimensiones son, simplemente, matrices $1 \times n$ y, por tanto, satisfacen todas sus propiedades algebraicas de suma y producto por escalar. Además:

a) Si $n = 1$, $\mathbb{R}^1$, que denotaremos simplemente mediante $\mathbb{R}$, representa gráficamente, ya sabemos (Volumen 0 (Fundamentos)), una *línea recta* (*o recta real*), donde se ha elegido un punto de referencia u origen al cual se le asigna el valor cero. En este espacio se acostumbra denotar los vectores sin el uso de los paréntesis; es decir, (x_1) se escribe simplemente como x_1. En la figura 5 tenemos la representación gráfica de $\mathbb{R}$.

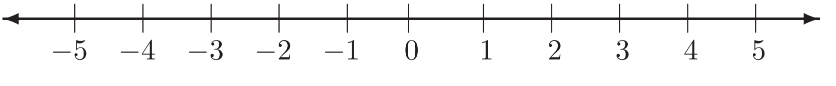


Figura 5. Recta real $\mathbb{R}$

b) Si $n = 2$, $\mathbb{R}^2$ representará el *plano bidimensional*. Con el fin de localizar un vector de $\mathbb{R}^2$, es conveniente elegir como sistema de referencia, ya sabemos (Volumen 0 (Fundamentos)), a dos rectas dirigidas que se cortan perpendicularmente en un punto al cual se le asigna el vector $(0,0)$. A este punto se le denomina *origen del sistema de coordenadas*. En la figura 6 vemos una representación del vector $(2,3)$ en este espacio.

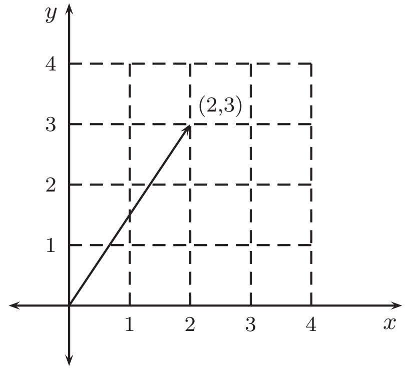


Figura 6. Plano bidimensional $\mathbb{R}^2$

c) Para $n = 3$, $\mathbb{R}^3$ representa el *espacio tridimensional*. El sistema de referencia en este espacio, también sabemos (Volumen 0 (Fundamentos)), está conformado por tres rectas dirigidas que se cortan de forma ortogonal en un punto que corresponde al origen del sistema de coordenadas. A este punto le asignamos el vector $(0,0,0)$. En la figura 7 hemos localizado los vectores $(1,1,1)$ y $(1,0,1)$.

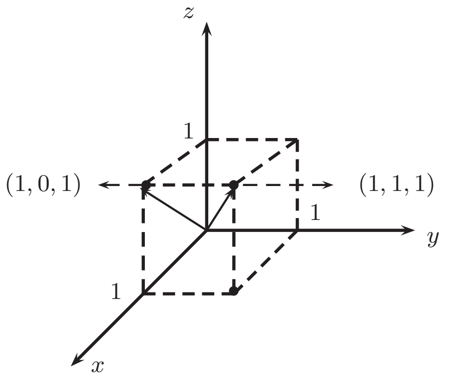


Figura 7. Espacio tridimensional $\mathbb{R}^3$

De acuerdo con la definición 1, no es muy difícil entonces saber cuándo dos vectores son iguales:

Definición 2. (Igualdad de vectores)
Sean x, $y \in \mathbb{R}^n$. Decimos que $x = (x_1, x_2, ..., x_n)$ y $y = (y_1, y_2, ..., y_n)$ son *iguales* si cada componente del vector x es igual a la correspondiente componente del vector y; es decir, si $x_i = y_i$, para $i = 1, 2, \ldots, n$.

Y aunque el álgebra de vectores está definida por la observación de que un vector es una matriz $1 \times n$ y, por lo tanto, hereda las propiedades matriciales que estudiamos en la lección 2, es conveniente reafirmarla una vez más, pero ahora en este contexto.

Definición 3. (Álgebra de vectores (Möbius (1827), Gibbs (1881))) Sean $x = (x_1, x_2, ..., x_n)$, $y = (y_1, y_2, ..., y_n) \in \mathbb{R}^n$ dos vectores cualesquiera. Entonces:

a) La *suma* $x + y$ de los vectores x y y es el vector cuyas componentes son la suma de las componentes correspondientes de x y y; es decir,

$$x + y = (\, x_1 + y_1,\, x_2 + y_2,\, \ldots,\, x_n + y_n \,)$$

b) La *resta o diferencia* de los vectores x y y es el vector cuyas componentes son la resta de las componentes correspondientes de x y y; es decir,

$$x - y = (\, x_1 - y_1,\, x_2 - y_2,\, \ldots,\, x_n - y_n \,)$$

c) Si $k \in \mathbb{R}$ entonces definimos el *producto por escalar* kx como el vector cuyas componentes son las componentes de x multiplicadas por k; es decir,

$$kx = (\, kx_1,\, kx_2,\, \ldots,\, kx_n \,)$$

En el caso de vectores con dos componentes, puede verificarse gráficamente el significado de la operación suma. Ésta resulta, precisamente, sobre la *diagonal del paralelogramo* generado por estos dos vectores, tal y como se muestra en la figura 8. Por esta razón, se dice que la suma de vectores satisface la *ley del paralelogramo*.

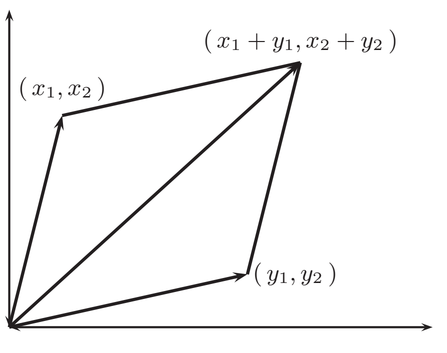


Figura 8. Suma de vectores o ley del paralelogramo

Podemos también representar geométricamente la diferencia entre los vectores x y y. Como $x = (x - y) + y$, entonces el vector $x - y$ es el vector que debe ser sumado a y para obtener x. Esto se representa en la figura 9.

La multiplicación por escalar podemos verla representada en la figura 10, donde se muestra al vector kx, para $k > 1$.

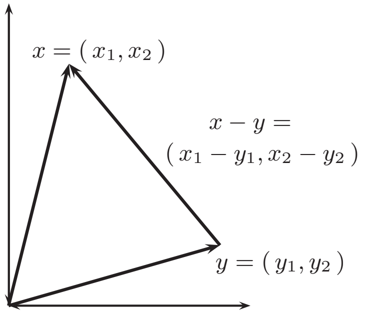


Figura 9. Diferencia de vectores

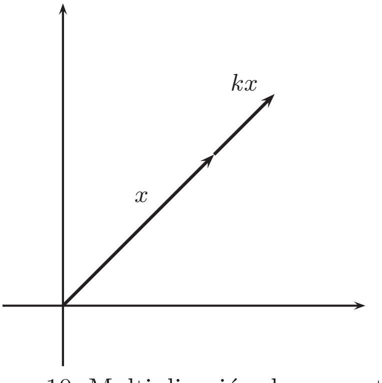


Figura 10. Multiplicación de un vector por un escalar

Ejemplo 1.
Para los vectores en $\mathbb{R}^2$, $x = (1, 2)$ y $y = (3, 5)$, encontremos $x + y$, $x - y$, $5x$ y $-2y$.

Solución

a) $x + y = (1, 2) + (3, 5) = (4, 7)$ b) $x - y = (1, 2) - (3, 5) = (-2, -3)$

c) $5x = 5(1, 2) = (5, 10)$ d) $-2y = -2(3, 5) = (-6, -10)$

Definición 4. (Vectores paralelos)
Dos vectores x y y en $\mathbb{R}^n$, diferentes de cero, son *paralelos* si existe $k \in \mathbb{R}$ tal que $x = ky$ (figura 11).

Ejemplo 2.

a) Los vectores $x = (1, 1)$ y $y = (2, 2)$ son paralelos porque $x = \frac{1}{2}y$.

b) Los vectores $x = (1, 1)$ y $y = (-1, -1)$ también son paralelos ya que $x = (-1)y$.

c) Los vectores $x = (-6, 12)$ y $y = (-2, 4)$ también son paralelos porque $x = 3y$.

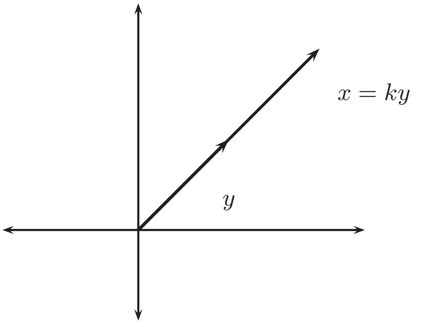
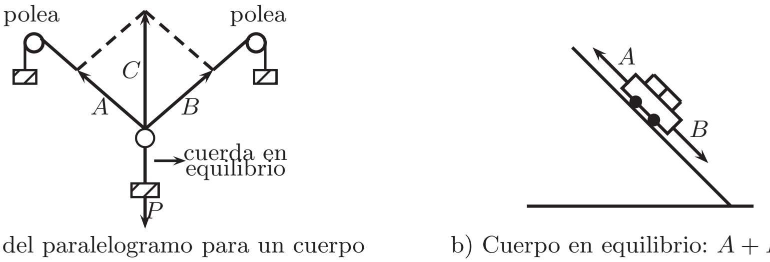


Figura 11. Vectores paralelos

Nota 2. (**Sobre la ley del paralelogramo**)
La *ley del paralelogramo* para la suma de vectores es tan intuitiva que sus orígenes parecen perdidos en la historia. Podría haber aparecido en uno de los trabajos de Aristóteles [384-322 a. C.] que se destruyeron, aunque la idea sí está descrita explícitamente en la *Mecánica* de Heron de Alejandría (siglo I d. C.), y también está en el primer corolario de los *Principia Mathematica* (1686) de Isaac Newton [1642-1727] aunque, en los *Principia*, Newton trata con lo que hoy podríamos considerar como entidades vectoriales, tales como fuerzas y velocidades, pero nunca con el concepto mismo de vector.

Nota 3.
En adelante, el lector no debería confundirse con el uso simultáneo de las palabras *punto* y *vector* para referirse, por ejemplo, al par $(x_1, x_2) \in \mathbb{R}^2$. Lo único que debería tenerse en cuenta es que, en algunas ocasiones, es más útil pensar las parejas de números como vectores que como puntos.

Nota 4. (**Sobre el concepto de equilibrio físico**)
Cabe anotar aquí que en mecánica se dice que un cuerpo está en *equilibrio físico* si la *suma* de las fuerzas que actúan sobre él suman cero. Para ilustrar esto véanse las figuras 12 a) y 12 b).



a) Ley del paralelogramo para un cuerpo en equilibrio: $P = C$ y $C = A + B$

b) Cuerpo en equilibrio: $A + B = 0$

Figura 12. Equilibrio físico

Ejercicios 1

1) Si $x = (1, 2)$, $y = (-1, 3)$, calcule

a) $x + y$	b) $x - y$
c) $3x + 2y$	d) $-x + 5y$
e) $-x - y$	f) $7x + 10y$

2) Si $x = (3, -1, 2)$, $y = (8, 1, -2)$, calcule

a) $x + y$	b) $y - x$
c) $5x - 4y$	d) $3x + 2y$
e) $8x - 3y$	f) $-x - y$

3) Dibuje en el plano $\mathbb{R}^2$ los vectores $x = (1, 2)$, $y = (-1, 3)$; y en el espacio $\mathbb{R}^3$ los vectores $x = (3, -1, 2)$, $y = (8, 1, -2)$.

4) Determine gráficamente el vector que va desde $x = (1, 1)$ a $y = (3, 4)$. Compruebe que este vector es $y - x$.

5) ¿Cuáles de los siguientes vectores son paralelos entre sí?:

a) $(1, 1)$	b) $(-1, -1)$	c) $(3, 5)$
d) $(7, 14)$	e) $(1, 2)$	f) $(50, 50)$

2. Norma de un vector en $\mathbb{R}^n$

La definición de norma o longitud de un vector (término, el primero, acuñado por Gauss en 1832) es muy natural para $n = 2$ y $n = 3$, pues se calcula en forma precisa mediante el teorema de Pitágoras de la geometría euclidiana considerando el vector únicamente como un segmento de recta, e ignorando sus otras características como la dirección y la magnitud. Sin embargo, este dato es fundamental cuando entendemos que esta longitud mediría en vectores físicos tales como fuerza o velocidad, sus correspondientes magnitudes. Veamos entonces esta definición central en la teoría de vectores.

Definición 5. (Norma de un vector y distancia en $\mathbb{R}^2$)
Para un vector $x = (x_1, x_2) \in \mathbb{R}^2$, definimos su *norma* (o *longitud*), la cual denotaremos por $\| x \|$, como

$$\| x \| = \sqrt{x_1^2 + x_2^2}$$

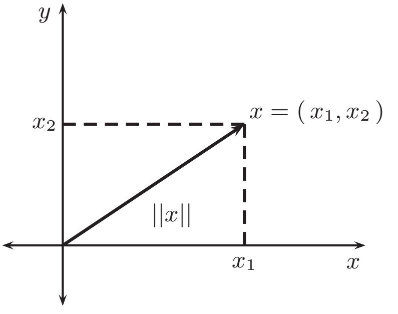


Figura 13. Norma en $\mathbb{R}^2$ de un vector = longitud del vector

Y utilizando la definición anterior, definimos la *distancia* entre dos puntos x y y de $\mathbb{R}^2$ como la longitud del vector $x - y$. Si $x = (x_1, x_2)$ y $y = (y_1, y_2)$, entonces la distancia entre x y y es, por el mismo teorema de Pitágoras,

$$\| x - y \| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

Definición 6. (Norma de un vector y distancia en $\mathbb{R}^3$)
Para un vector $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ definimos su *norma*, la cual denotaremos por $\| x \|$, como

$$\| x \| = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

Y la *distancia* entre $x = (x_1, x_2, x_3)$ y $y = (y_1, y_2, y_3)$ es la longitud del vector $x - y$; es decir,

$$\| x - y \| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}$$

Sin embargo, muchos problemas matemáticos requieren una generalización para $n \geq 4$ de los conceptos de longitud de un vector y de distancia entre vectores. Es "natural" definirlos así:

Definición 7. (Norma de un vector y distancia en $\mathbb{R}^n$)
Para un vector $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ definimos su *norma* como

$$\| x \| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

y la *distancia* entre $x = (x_1, x_2, \ldots, x_n)$ y $y = (y_1, y_2, \ldots, y_n)$ se define como la longitud del vector $x - y$; es decir,

$$\| x - y \| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}$$

Definición 8. (Vector unitario o de norma 1)
Un vector $x \in \mathbb{R}^n$ se llama *unitario* (o de norma 1) si, y sólo si, $\| x \| = 1$ (figura 14).

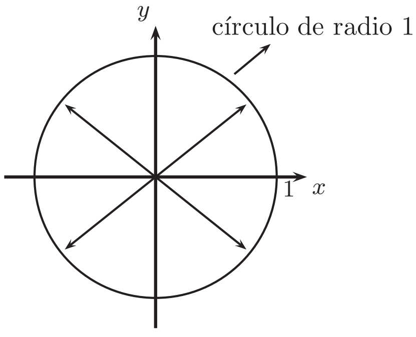


Figura 14. Vectores unitarios en $\mathbb{R}^2$

Ejemplo 3.

a) La norma de $x = (3,4) \in \mathbb{R}^2$ es $|| x || = \sqrt{3^2 + 4^2} = 5$

b) La distancia entre los puntos $x = (2,3)$ y $y = (4,-5)$ en $\mathbb{R}^2$ es

$$\| x - y \| = \sqrt{(2-4)^2 + (3-(-5))^2} = \sqrt{(-2)^2 + 8^2} = \sqrt{68}$$

c) La norma del vector $x = (1, -2, 4) \in \mathbb{R}^3$ es

$$\| x \| = \sqrt{1^2 + (-2)^2 + 4^2} = \sqrt{21}$$

d) La distancia entre los puntos $x = (1, -2, 4)$ y $y = (0, 1, 2)$ en $\mathbb{R}^3$ es

$$\| x - y \| = \sqrt{(1-0)^2 + (-2-1)^2 + (4-2)^2} = \sqrt{14}$$

e) Los vectores $(1,0)$ y $(0,1)$ son unitarios en $\mathbb{R}^2$.

f) Los vectores $i \equiv (1,0,0)$, $j \equiv (0,1,0)$, $k \equiv (0,0,1)$ son unitarios en $\mathbb{R}^3$.

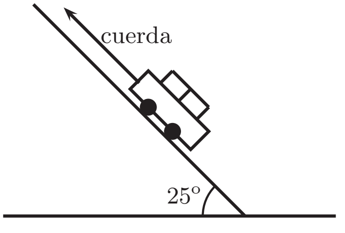
g) Los vectores $\left(\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}} \right)$ y $\left(\dfrac{2}{\sqrt{5}}, \dfrac{1}{\sqrt{5}} \right)$ son unitarios en $\mathbb{R}^2$.

h) En general, si $x \in \mathbb{R}^n$, $x \neq 0$, entonces $\dfrac{x}{\| x \|}$ es siempre un vector unitario en $\mathbb{R}^n$.

Ejemplo 4. (Componente de una fuerza en una dirección dada)
Determinemos qué fuerza en la cuerda de la figura 15 mantendrá en *equilibrio físico* un automóvil de 5,000 libras si la rampa forma un ángulo de 25º con la horizontal.

Solución

Al introducir coordenadas, el peso del automóvil es $\vec{p} = [\,0, -5000\,]$ ya que esta fuerza apunta hacia abajo en la dirección del eje Y negativo (figura 15).



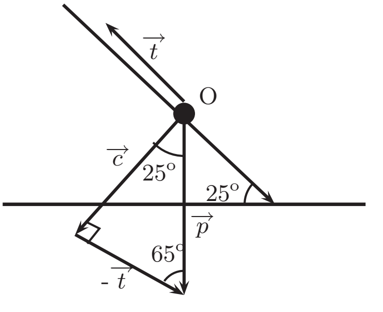


Figura 15

Aquí, es claro que $\vec{p} = \vec{c} - \vec{t}$, donde $\vec{c}$ es la fuerza perpendicular que ejerce el automóvil sobre la rampa, y $\vec{t}$ es la tensión de la cuerda que tiene magnitud

$$\|\,\vec{t}\,\| = \|\,\vec{p}\,\| \cos 65^{\text{o}} = 5{,}000 \cos 65^{\text{o}} = 2{,}113 \text{ libras}$$

Luego la fuerza $\vec{t}$ de equilibrio es (figura 16)

$$\vec{t} = 2{,}113(\,-\cos 25^{\text{o}}, \operatorname{sen} 25^{\text{o}}\,) = (\,-1{,}915.03\ ,\ 892.99\,)$$

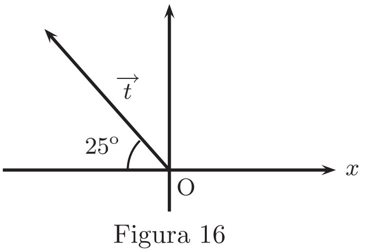


Figura 16

Ejercicios 2

1) Calcule la norma (longitud) de los siguientes vectores:

a) $(-2, 0)$	b) $(3, 1)$
c) $(1, 7)$	d) $(-4, -2)$
e) $(1, 7, 2)$	f) $(-1, 4, 9)$

Dibuje cada uno de ellos.

2) Calcule la distancia entre los siguientes puntos:

a) $(5, 1)$ y $(3, 4)$	b) $(-2, 2)$ y $(1, 1)$
c) $(2, 1, 3)$ y $(3, 2, 4)$	d) $(4, -2, 2)$ y $(1, 1, 6)$
e) $(-1, 0, 7)$ y $(-1, 2, 4)$	f) $(1, 1, 1)$ y $(2, 2, 2)$

3) Verifique que el triángulo cuyos vértices son $(1, 1)$, $(5, 4)$ y $(-2, 5)$ es isósceles (es decir, que dos de sus lados son iguales) (Volumen 0 (Fundamentos)).

4) Verifique que el cuadrilátero cuyos vértices son $(-6, -2)$, $(-2, -1)$, $(-1, 3)$ y $(-5, 2)$ es un rombo (es decir, un paralelogramo con sus cuatro lados iguales) (Volumen 0 (Fundamentos))

5) Verifique que el triángulo cuyos vértices son $(-1, -3)$, $(2, -1)$ y $(-2, 5)$ es rectángulo.

6) Verifique que los puntos $(-5, 7)$, $(2, 6)$ y $(1, -1)$ están sobre una circunferencia de radio 5 y centro en $(-2, 3)$.

3. Ángulo entre vectores

El concepto originalmente geométrico de ángulo tiene su expresión dentro de la geometría analítica en la noción de ángulo entre dos vectores que, siguiendo el pensamiento de Descartes y Fermat, y ya no el de los griegos antiguos, ahora se puede introducir de forma natural bajo las siguientes definiciones y consideraciones.

Definición 9. (Producto interior (Clifford (1878)))
El *producto interior (o producto punto)* entre los vectores $x = (x_1, x_2, \ldots, x_n)$ y $y = (y_1, y_2, \ldots, y_n)$ de $\mathbb{R}^n$ es el valor

$$x \cdot y = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \sum_{i=1}^{n} x_iy_i$$

Nota 5.
Observemos que ésta no es, realmente, una nueva definición. Si identificamos un vector $x \in \mathbb{R}^n$ como una matriz $1 \times n$, entonces el producto interior de x con otro vector $y \in \mathbb{R}^n$ es, simplemente,

$$x \cdot y = xy^T$$

donde y^T es la matriz $n \times 1$ traspuesta de la matriz y, y el producto de la parte derecha de la igualdad es el producto usual entre matrices.

Ejemplo 5.

a) $(2, 3, 1) \cdot (4, -5, 2) = (2)(4) + (3)(-5) + (1)(2) = -5$

b) $(4, 5) \cdot (3, -3) = (4)(3) + (5)(-3) = -3$

c) $(3, 1, 6, 2) \cdot (4, 2, 1, 9) = (3)(4) + (1)(2) + (6)(1) + (2)(9) = 38$

d) $(3, 12, 4, 1, -7) \cdot (8, 1, -6, 3, -9) = (3)(8) + (12)(1) + (4)(-6) + (1)(3) + (-7)(-9) = 78$ ▲

Y aunque aquí podríamos recurrir a la Nota 5 anterior para deducir las propiedades algebraicas del producto interno, a partir de las propiedades de la multiplicación de matrices y de la definición de matriz traspuesta, preferimos abordar estas propiedades desde la definición inmediata. Veamos esto entonces.

Teorema 1. ***(Propiedades del producto interior)***
Sean x, y, z vectores en $\mathbb{R}^n$ y k un número real. Entonces

a) $x \cdot y = y \cdot x$

b) $x \cdot (y + z) = x \cdot y + x \cdot z$

c) $x \cdot (ky) = (kx) \cdot y = k(x \cdot y)$

d) $\| x \| = \sqrt{x \cdot x}$

e) $x \cdot x = 0$ *si, y sólo si,* $x = 0$

Demostración
Sean $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, $y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$. Entonces

a) $x \cdot y = \sum_{i=n}^{n} x_i y_i = \sum_{i=n}^{n} y_i x_i = y \cdot x$

b) $x \cdot (y + z) = \sum_{i=n}^{n} x_i(y_i + z_i) = \sum_{i=n}^{n} x_i y_i + \sum_{i=n}^{n} x_i z_i = x \cdot y + x \cdot z$

c) $x \cdot (\alpha y) = \sum_{i=n}^{n} x_i(\alpha y_i) = \sum_{i=n}^{n} (\alpha x_i) y_i = \alpha \sum_{i=n}^{n} x_i y_i = \alpha(x \cdot y)$

d) $\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{x \cdot x}$

e) Es claro que si $x = 0$, entonces $x \cdot x = 0$. De otro lado, si $x \cdot x = 0$, entonces $\sum_{i=n}^{n} x_i^2 = 0$, lo que implica que $x_i = 0$ para todo $i = 1, \ldots, n$. ■

Ejemplo 6.
Sean $x = (3, 4, 1)$, $y = (5, -2, 2)$, $z = (6, 1, 6)$. Entonces $x \cdot y = 15 - 8 + 2 = 9$, $x \cdot z = 18 + 4 + 6 = 28$, $y + z = (11, -1, 8)$, $3x = (9, 12, 3)$. Y observemos que $x \cdot (y + z) = (3, 4, 1) \cdot (11, -1, 8) = 33 - 4 + 8 = 37 = x \cdot y + x \cdot z$, $(3x) \cdot y = (9, 12, 3) \cdot (5, -2, 2) = 45 - 24 + 6 = 27 = 3(x \cdot y)$ ▲

Ahora: con el fin de encontrar una expresión para el *ángulo* θ que forman dos vectores x, $y \in \mathbb{R}^n$, observemos que

$$\| x - y \|^2 = (x - y) \cdot (x - y) = x \cdot x - 2x \cdot y + y \cdot y = \| x \|^2 + \| y \|^2 - 2x \cdot y \qquad (1)$$

Esta última igualdad está estrechamente relacionada con la conocida *ley de los cosenos* (Volumen 0 (Fundamentos)). Para entender por qué, consideremos la figura 17.

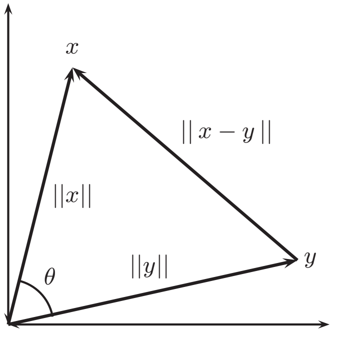


Figura 17. Ley de los cosenos

De la geometría griega sabemos que para el triángulo de la figura 17 debemos tener

$$\| x - y \|^2 = \| x \|^2 + \| y \|^2 - 2 \| x \| \| y \| \cos\theta \qquad \textit{(ley de los cosenos)}$$

Comparando esto con la igualdad (1) vemos entonces que necesariamente

$$x \cdot y = \| x \| \, \| y \| \cos \theta$$

Y esto induce la siguiente definición general:

Definición 10. (Ángulo entre vectores)
El *ángulo* entre los vectores $x, y \in \mathbb{R}^n$, $x \neq 0, y \neq 0$, lo definiremos como

$$\theta = \arccos \frac{x \cdot y}{\| x \| \, \| y \|}, \qquad 0 \leq \theta \leq \pi$$

Así, $x \cdot y > 0$ si, y sólo si, $0 < \theta < \pi/2$; $x \cdot y < 0$ si, y sólo si $\frac{\pi}{2} < \theta < \pi$.

Ejemplo 7.

a) Calculemos el ángulo θ formado por los vectores $x = (1, 4)$ y $y = (5, 3)$.

b) Calculemos el ángulo θ formado por los vectores $x = (0, 1, 1)$ y $y = (1, 1, 1)$.

c) Calculemos el ángulo θ formado por los vectores $x = (3, 1, 4, 5)$ y $y = (1, -2, 2, 6)$.

Solución

a) Aquí, $|| x || = \sqrt{1^2 + 4^2} = \sqrt{17}$; $|| y || = \sqrt{5^2 + 3^2} = \sqrt{34}$ y $x \cdot y = (1)(5) + (4)(3) = 17$. Por tanto, $\cos \theta = \dfrac{17}{\sqrt{17}\sqrt{34}} \approx 0.7071$. Luego, $\theta = 45^{\text{o}}$ o, equivalentemente, $\theta = 0.7853$ radianes.

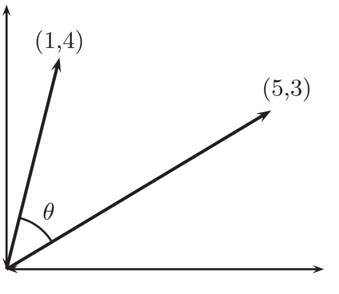


Figura 18

b) El ángulo θ podemos calcularlo utilizando la ecuación $2 = \sqrt{2}\sqrt{3} \cos \theta$. Por tanto, $\cos \theta = \sqrt{\frac{2}{3}}$. Luego, $\theta \cong 35.26^{\text{o}}$ o, equivalentemente, $\theta = 0.6154$ radianes.

c) Aquí,
$||\, x \,|| = \sqrt{3^2 + 1^2 + 4^2 + 5^2} = \sqrt{51}$; $||\, y \,|| = \sqrt{1^2 + (\,-2\,)^2 + 2^2 + 6^2}$
$= \sqrt{45}$ y $x \cdot y = (\,3\,)(\,1\,) + (\,1\,)(\,-2\,)(\,4\,)(\,2\,) + (\,5\,)(\,6\,) = 39.$
Por tanto, $\cos\theta = \frac{39}{\sqrt{51}\sqrt{45}} \approx 0.81409$. Luego, $\theta = 35.50^{\mathrm{o}}$ o, en radianes, $\theta = 0.6196$.

Ejemplo 8. (Trabajo realizado por una fuerza sobre un objeto)
Supongamos que tenemos un objeto sólido sobre el que actúa una fuerza constante $\vec{a}$, que le aplica un desplazamiento $\vec{d}$. El *trabajo* realizado por $\vec{a}$ en el desplazamiento se define como

$$T = \vec{a} \cdot \vec{d} = \|\,\vec{a}\,\|\,\|\,\vec{d}\,\|\cos\alpha$$

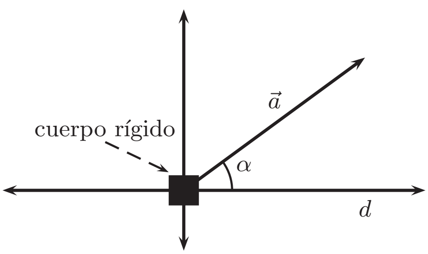


Figura 19

Así, si $\alpha = 90^{\mathrm{o}}$ entonces $T = 0$; pero si $\alpha > 90^{\mathrm{o}}$ entonces $T < 0$ (pues $\cos\alpha < 0$), y esto significa que en el desplazamiento existe trabajo *en contra* de la fuerza $\vec{a}$. ▲

Continuamos nuestro estudio con la que es, quizás, la desigualdad más importante de la teoría básica de vectores: *la desigualdad Cauchy-Schwarz*. Esta, como entenderemos más adelante, no es más que la antesala de la versión cartesiana del famoso teorema geométrico griego que afirma que en un triángulo cualquiera, la suma de las longitudes de cualquiera dos de sus lados es mayor o igual que la longitud del tercer lado (Volumen 0 (Fundamentos)).

Teorema 2. ***(Desigualdad Cauchy-Schwarz)***
Si x, y son dos vectores en $\mathbb{R}^n$, entonces

$$|\, x \cdot y \,| \leq ||\, x \,||\, ||\, y \,||$$

Demostración

Observemos que

$$
\begin{aligned}
0 \leq \left(\frac{x}{||\, x\, ||} - \frac{y}{||\, y\, ||} \right) \cdot \left(\frac{x}{||\, x\, ||} - \frac{y}{||\, y\, ||} \right) &= \frac{x \cdot x}{||\, x\, ||^2} - \frac{2x \cdot y}{||\, x\, ||\, ||\, y\, ||} + \frac{y \cdot y}{||\, y\, ||^2} \\
&= 1 - \frac{2x \cdot y}{||\, x\, ||\, ||\, y\, ||} + 1 \\
&= 2 - \frac{2x \cdot y}{||\, x\, ||\, ||\, y\, ||}
\end{aligned}
$$

Por tanto, $x \cdot y \leq ||\, x\, ||\ ||\, y\, ||$. Además,

$$
\begin{aligned}
0 \leq \left(\frac{x}{||\, x\, ||} + \frac{y}{||\, y\, ||} \right) \cdot \left(\frac{x}{||\, x\, ||} + \frac{y}{||\, y\, ||} \right) &= \frac{x \cdot x}{||\, x\, ||^2} + \frac{2x \cdot y}{||\, x\, ||\ ||\, y\, ||} + \frac{y \cdot y}{||\, y\, ||^2} \\
&= 1 + \frac{2x \cdot y}{||\, x\, ||\ ||\, y\, ||} + 1 \\
&= 2 + \frac{2x \cdot y}{||\, x\, ||\ ||\, y\, ||}
\end{aligned}
$$

Por tanto, también $||\, x\, ||\ ||\, y\, || \geq -x \cdot y$. De las desigualdades $x \cdot y \leq ||\, x\, ||\ ||\, y\, ||$ y $-||\, x\, ||\ ||\, y\, || \leq x \cdot y$ se tiene que

$$|\, x \cdot y\, | \leq ||\, x\, ||\ ||\, y\, || \quad \blacksquare$$

Nota 6.

La desigualdad de Cauchy-Schwarz también se conoce como "desigualdad de Buniakovsky" o "desigualdad de Kantorovich". Curiosamente, el nombre "Cauchy-Schwarz" equivoca la escritura del segundo apellido pues la desigualdad tiene este nombre por el matemático Hermann Amandus Schwartz [1843-1921].

Definición 11. (Vectores ortogonales)

Dos vectores x y y distintos de cero en $\mathbb{R}^n$ son *ortogonales (o perpendiculares)* si el ángulo formado por ellos es $\dfrac{\pi}{2}$ (figura 20).

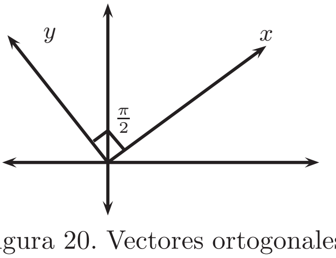


Figura 20. Vectores ortogonales

Ejemplo 9. (Vectores ortogonales en $\mathbb{R}^2$ y en $\mathbb{R}^3$)

a) $x = (1, 2)$, $y = (-2, 1)$ son ortogonales en $\mathbb{R}^2$

b) $x = (-1, 3, 2)$, $y = (2, 4, -5)$ son ortogonales en $\mathbb{R}^3$

Teorema 3. ***(Propiedades de la norma de un vector)***
Si x, y son dos vectores en $\mathbb{R}^n$, $k \in \mathbb{R}$, entonces

a) Si $x,\ y \neq 0,\ x \cdot y = 0$ si, y sólo si, x es ortogonal a y

b) $||\, kx \,|| = |\, k \,|\, ||\, x \,||$

c) $||\, x + y \,|| \leq ||\, x \,|| + ||\, y \,||$ *(desigualdad triangular)*

d) $||\, x + y \,||^2 + ||\, x - y \,||^2 = 2(||\, x \,||^2 + ||\, y \,||^2)$

Demostración

a) i) Sean x y y dos vectores distintos de cero. Si los vectores son ortogonales, entonces, por el teorema de Pitágoras, $||\, x - y \,||^2 = ||\, x \,||^2 + ||\, y \,||^2$; es decir, $x \cdot x - 2x \cdot y + y \cdot y = x \cdot x + y \cdot y$. Por tanto, $x \cdot y = 0$.

ii) Supongamos ahora que $x \cdot y = 0$. Esto implica que

$$||\, x - y \,||^2 = ||\, x \,||^2 + ||\, y \,||^2$$

Utilizando la ley de los cosenos (figura 17) se tiene que

$$||\, x \,||^2 + ||\, y \,||^2 - 2||\, x \,||\ ||\, y \,|| \cos\theta = ||\, x \,||^2 + ||\, y \,||^2$$

Como x y y son distintos de cero, entonces $\cos\theta = 0$; luego $\theta = \frac{\pi}{2}$.

b) $||\, kx \,|| = \sqrt{(\, k \cdot x\,) \cdot (\, k \cdot x\,)} = \sqrt{k^2(\, x \cdot x\,)} = |\, k \,|\, ||\, x \,||$

c) Observemos que

$$\begin{aligned} ||\, x + y \,||^2 &= (\, x + y\,) \cdot (\, x + y\,) = x \cdot x + 2x \cdot y + y \cdot y \\ &= ||\, x \,||^2 + 2x \cdot y + ||\, y \,||^2 \ \leq ||\, x \,||^2 + 2||\, x \,|| \cdot ||\, y \,|| + ||\, y \,||^2 \\ &= (\, ||\, x \,|| + ||\, y \,||\,)^2 \end{aligned}$$

Por tanto, $||\, x + y \,|| \leq ||\, x \,|| + ||\, y \,||$. (¿En qué punto de la demostración se aplicó la desigualdad Cauchy-Schwarz?)

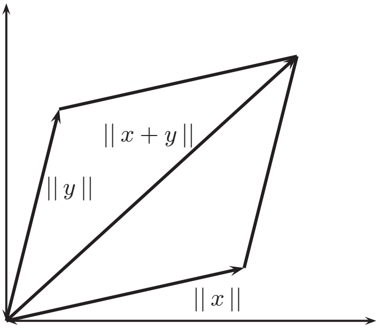


Figura 21. Desigualdad triangular

d) Basta observar que $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2x \cdot y$ y que $\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2x \cdot y$; así $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y^2\|)$ (¿Qué resultado geométrico describe esta igualdad?) ■

Ejemplo 10.

a) Si $x = (1, 3)$, $y = (-1, -2)$ entonces $|x \cdot y| = 7$, $\|x\|\|y\| = \sqrt{10}\sqrt{5} = \sqrt{50} \cong 7.07$; así $|x \cdot y| \leq \|x\|\|y\|$, ilustrando la desigualdad Cauchy-Schwarz.

b) Y con los mismos vectores de a), $\|x + y\| = \|(0, 1)\| = 1 \leq \|x\| + \|y\| = \sqrt{10} + \sqrt{5} \cong 5.39$ corroborando la desigualdad triangular.

a. Proyección de un vector sobre otro

Sean x y y dos vectores en $\mathbb{R}^n$ con $y \neq 0$, y supongamos que es posible encontrar un número c tal que $x - cy$ sea perpendicular a y. ¿Cuál es este número c? (figura 22).

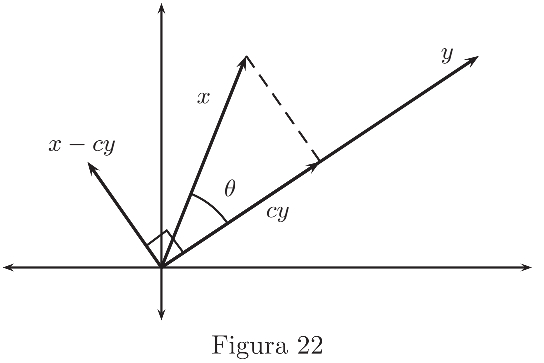


Figura 22

Como $x - cy$ es perpendicular a y, entonces $(\, x - cy\,) \cdot y = 0$ y, por tanto,

$$c = \frac{x \cdot y}{||\, y\, ||^2}$$

Y esto da origen a la siguiente definición que, como siempre aquí, está inspirada en la geometría clásica:

Definición 12. (Proyección de x sobre y)
Sean x y $y \in \mathbb{R}^n$, $y \neq 0$. Definimos la *proyección de x a lo largo de y* como el vector

$$\frac{x \cdot y}{||\, y \,||^2} y$$

Ejemplo 11.
Sean a) $x = (\,3, 2\,)$ y $y = (\,4, 6\,)$; b) $x = (\,1, -2, 3\,)$ y $y = (\,4, 1, 5\,)$. Calculemos la proyección de x a lo largo de y.

Solución

a) Como $x \cdot y = 24$ y $||\, y \,||^2 = 4^2 + 6^2 = 52$, entonces la proyección de x a lo largo de y es el vector

$$\frac{24}{52}(\,4, 6\,) = \left(\frac{24}{13}, \frac{36}{13}\right)$$

b) Como $x \cdot y = 17$ y $||\, y \,||^2 = 4^2 + 1^2 + 5^2 = 42$, entonces la proyección de x a lo largo de y es el vector

$$\frac{17}{42}(\,4, 1, 5\,) = \left(\frac{34}{21}, \frac{17}{42}, \frac{85}{42}\right)$$

b. Producto cruz de vectores

La multiplicación interior nos provee de un escalar como producto de dos vectores. En su lugar, la *multiplicación cruz* de ellos nos muestra otro *vector* que tendrá propiedades geométricas y físicas importantes en el espacio tridimensional y que, además, es fácil de calcular.

Definición 13. (Producto cruz (Clifford (1878)))
El *producto cruz (o producto vectorial)* entre los vectores $x = (\,x_1, x_2, x_3\,)$ y $y = (\,y_1, y_2, y_3\,)$ de $\mathbb{R}^3$ es el vector, denotado como $x \times y$, dado por

$$x \times y = (\,x_2y_3 - x_3y_2,\; x_3y_1 - x_1y_3,\; x_1y_2 - x_2y_1\,)$$

o, abusando un tanto de la notación de determinantes,

$$x \times y = \begin{vmatrix} i & j & k \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}$$

donde, en notación del propio Clifford, definimos $i \equiv (1, 0, 0)$, $j \equiv (0, 1, 0)$, $k \equiv (\,0, 0, 1\,)$.

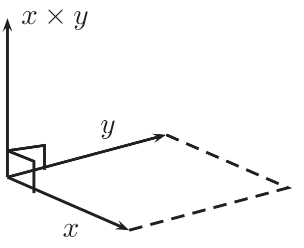


Figura 23. Producto cruz $x \times y$

La razón de esta definición aparentemente extraña, comienza a dilucidarse cuando observamos que $x \times y$ es *siempre ortogonal tanto a x como a y* (figura 23). En efecto, notemos que

$$\begin{aligned} x \cdot (x \times y) &= (x_1, x_2, x_3)(x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1) \\ &= x_1x_2y_3 - x_1x_3y_2 + x_2x_3y_1 - x_2x_1y_3 + x_3x_1y_2 - x_3x_2y_1 \\ &= 0 \end{aligned} \tag{1}$$

y además que

$$\begin{aligned} y \cdot (x \times y) &= (y_1, y_2, y_3)(x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1) \\ &= y_1x_2y_3 - y_1x_3y_2 + y_2x_3y_1 - y_2x_1y_3 + y_3x_1y_2 - y_3x_2y_1 \\ &= 0 \end{aligned} \tag{2}$$

Más aún: si x y y son paralelos, entonces $x \times y = 0$, pues si $x = cy$ para $c \in \mathbb{R}$, entonces

$$x \times y = \begin{vmatrix} i & j & k \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = \begin{vmatrix} i & j & k \\ cy_1 & cy_2 & cy_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = c \begin{vmatrix} i & j & k \\ y_1 & y_2 & y_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = 0$$

pues este último determinante tiene dos filas iguales.

Ejemplo 12.
Calculemos el producto cruz de los siguientes vectores:

a) $x = (2, 3, 1)$ y $y = (4, -5, 2)$ b) $x = (7, -2, -3)$ y $y = (2, 1, 1)$

Solución

a) $x \times y = (2, 3, 1) \times (4, -5, 2) = (3(2) - 1(-5), 1(4) - 2(2), 2(-5) - 3(4)) = (11, 0, -22)$

b) $x \times y = (7, -2, -3) \times (2, 1, 1) = ((-2)1 - (-3)1, (-3)2 - 7(1), 7(1) - (-2)2) = (1, -13, 11)$ ▲

Las propiedades fundamentales de esta nueva operación vectorial son las siguientes:

Teorema 4. ***(Propiedades del producto cruz)***
Sean x, y, $z \in \mathbb{R}^3$. Entonces

a) $x \times y = -(y \times x)$ *(no-conmutatividad)*

b) $x \cdot (x \times y) = 0$, $y \cdot (x \times y) = 0$

c) $x \times (y + z) = (x \times y) + (x \times z)$

d) $k(x \times y) = (kx) \times y = x \times (ky)$, *para todo* $k \in \mathbb{R}$

e) $\|x \times y\| = \|x\| \, \|y\| \operatorname{sen} \theta$, *donde* θ *es el ángulo formado por los vectores* x, y.

Demostración

a) Sean $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3)$; entonces

$$x \times y = \begin{vmatrix} i & j & k \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = - \begin{vmatrix} i & j & k \\ y_1 & y_2 & y_3 \\ x_1 & x_2 & x_3 \end{vmatrix} = -(y \times x)$$

b) Esta condición fue probada en las igualdades (1) y (2) al principio de esta sección.

c) Sean $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3)$, $z = (z_1, z_2, z_3)$; entonces

$$x \times (y + z) = \begin{vmatrix} i & j & k \\ x_1 & x_2 & x_3 \\ y_1 + z_1 & y_2 + z_2 & y_3 + z_3 \end{vmatrix}$$

$$= \begin{vmatrix} i & j & k \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} + \begin{vmatrix} i & j & k \\ x_1 & x_2 & x_3 \\ z_1 & z_2 & z_3 \end{vmatrix}$$

$$= (x \times y) + (x \times z)$$

d) Sean $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3)$, y $c \in \mathbb{R}$; entonces

$$c(x \times y) = c \begin{vmatrix} i & j & k \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = \begin{vmatrix} i & j & k \\ cx_1 & cx_2 & cx_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = (cx) \times y$$

$$c(x \times y) = c \begin{vmatrix} i & j & k \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = \begin{vmatrix} i & j & k \\ x_1 & x_2 & x_3 \\ cy_1 & cy_2 & cy_3 \end{vmatrix} = x \times (cy)$$

e) [Indicación: pruebe primero que $\|x \times y\|^2 = \|x\|^2 \|y\|^2 - (x \cdot y)^2$ y, por tanto, $\|x \times y\|^2 = \|x\|^2 \|y\|^2 (1 - \cos^2 \theta)$. El resultado es, de aquí, inmediato].

Nota 7.
Es posible interpretar geométricamente la parte e) del teorema 4, pues a partir de la figura 24, es fácil ver que el área del paralelogramo formado por x y y es igual a $\|x\| \; \|y\| \operatorname{sen} \theta$.

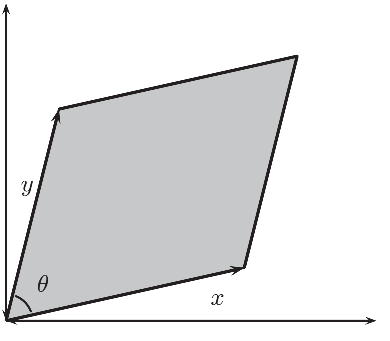


Figura 24

Ejemplo 13.
Calculemos el área del paralelogramo cuyos vértices adyacentes son $(-2, 1, 0)$, $(1, 4, 2)$ y $(-3, 1, 5)$.

Solución
El área del paralelogramo está dada por $\|x \times y\|$, donde $x = (3, 3, 2) = (1, 4, 2) - (-2, 1, 0)$ y $y = (-4, -3, 3) = (-3, 1, 5) - (1, 4, 2)$ (¿Por qué es necesario tomar estas diferencias?); es decir,

$$\| x \times y \| = \begin{vmatrix} i & j & k \\ 3 & 3 & 2 \\ -4 & -3 & 3 \end{vmatrix} = || (9 + 6, -8 - 9, -9 + 12) ||$$

$$= || (15, -17, 3) || = \sqrt{523}$$

Ejemplo 14. (Álgebra básica del producto cruz)
Es inmediato probar, a partir de la definición, que

$$i \times j = k, \quad j \times k = i, \quad k \times i = j, \quad j \times i = -k, \quad k \times j = -i, \quad i \times k = -j$$

Ejemplo 15.
Utilizando las seis igualdades del ejemplo 14 y el teorema 4 (*Propiedades del producto cruz*) es fácil calcular $x \times y$ para cualquier $x,\ y \in \mathbb{R}^3$. Ilustremos esto con un ejemplo:

Si $x = (3, 1, 2)$, $y = (1, 2, 9)$, entonces $x = 3i + j + 2k$, $y = i + 2j + 9k$, y así, utilizando la propiedad c) del teorema 4 anterior y las propiedades del ejemplo 14, se tendrá que:

$$\begin{aligned} x \times y &= (3i + j + 2k) \times (i + 2j + 9k) \\ &= 3i \times i + 6i \times j + 27i \times k + j \times i + 2j \times j + 9j \times k + 2k \times i + \\ &\quad 4k \times j + 18k \times k \\ &= (3)(0) + 6k - 27j - k + 2 \cdot 0 + 9i + 2j - 4i + 18 \cdot 0 \\ &= 5i - 25j + 5k \\ &= (5,\ -25,\ 5) \end{aligned}$$

Observemos que, efectivamente, $(3, 1, 2) \cdot (5, -25, 5) = 0$ y $(1, 2, 9) \cdot (5, -25, 5) = 0$.

Ejemplo 16. (Momento de una fuerza)
En mecánica, el *momento* m *de una fuerza* p *alrededor de un punto* Q se define como el producto $m = \| p \| d$, donde d es la distancia (perpendicular) entre Q y la línea de acción L de p (figura 25).

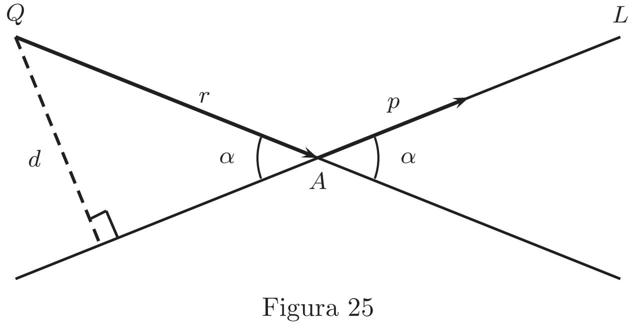


Figura 25

Si r es el vector que va de Q a cualquier punto A de L, entonces $d = \| r \| \operatorname{sen} \alpha$ y así $m = \| r \| \| p \| \operatorname{sen} \alpha$. Y como α es el ángulo entre r y p, entonces $m = \| r \times p \|$. Al vector $m = \| r \times p \|$ se le llama el *vector momento* de p alrededor de Q. Su magnitud es m y su dirección es la del eje de rotación alrededor de Q que produce p.

Nota 8. (Nota sobre el producto cruz)
El vector producto cruz *no* tiene las mismas propiedades de un vector ordinario; en cierta forma, es un *vector artificial.* Por ejemplo, no existe un vector producto cruz en el espacio de más de tres vectores. Es, sin duda, una herramienta útil de la mecánica y de la física pero sólo en el espacio $\mathbb{R}^3$.

Ejercicios 3

1) Calcule el producto interior $x \cdot y$ de los siguientes vectores:

   a) $x = (1, 2)$ y $y = (3, -1)$
   b) $x = (-7, -2)$ y $y = (1, 1)$
   c) $x = (3, 4, 2)$ y $y = (-1, -1, 1)$
   d) $x = (4, 1, 2, 0)$ y $y = (-3, 5, 9, 1)$

2) ¿Si $x \cdot y = x \cdot z$, entonces $y = z$?

3) Determine el ángulo formado por los vectores x y y del ejercicio 1) anterior.

4) Calcule α tal que $x = (3, \alpha, 1)$ sea ortogonal a $y = (1, -3, 7)$.

5) Verifique la desigualdad del triángulo $||\, x + y \,|| \leq ||\, x \,|| + ||\, y \,||$ en cada uno de los literales del ejercicio 1) anterior.

6) Encuentre el área de los siguientes paralelogramos cuyos vértices adyacentes son:

   a) $(3, 6, -1), (1, 1, 6), (-4, 2, 3)$ b) $(-1, 2, 5), (2, 2, 3), (-3, 3, 7)$

7) Encuentre la proyección de x sobre y en los siguientes casos (dibuje los casos en $\mathbb{R}^2$):

   a) $x = (1, 1)$, $y = (1, 0)$ b) $x = (3, 4)$, $y = (1, 1)$
   c) $x = (1, -1, 2)$, $y = (2, 1, 5)$ d) $x = (-5, 2, 7)$, $y = (3, 1, 4)$

8) Determine una fuerza P tal que $P = A - B$, donde $||\, A \,|| = 1$, $||\, B \,|| = 2$ y los ángulos con el eje x están dados en la figura 26. ¿Qué tipo de problema físico se está representando aquí?

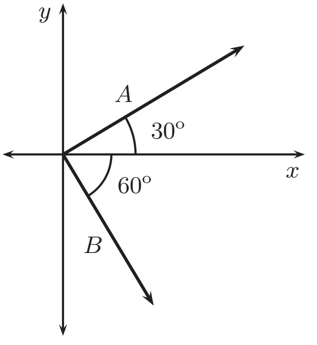


Figura 26

9) Calcule el producto cruz $x \times y$ de los siguientes vectores:

a) $x = (2, 2, 3)$, $y = (4, -3, -1)$

b) $x = (-7, -2, 1)$, $y = (2, 3, 5)$

4. Rectas y planos

La geometría analítica del plano y, en mayor dimensión, la del espacio (como también la teoría de las líneas rectas y de los planos) utilizan exclusivamente el aparato del álgebra lineal en su forma más simple: una línea recta en el plano se describe mediante una ecuación lineal en dos variables ligando sus dos coordenadas con un punto de la recta; un plano en el espacio está dado por una ecuación lineal en tres variables (que son sus coordenadas en el espacio) y cada tres coordenadas de esta ecuación coinciden con un punto del espacio. Debe anotarse, sin embargo, que ni Descartes ni Fermat desarrollaron la geometría analítica del espacio; sólo la del plano. Este trabajo fue comenzado posteriormente, en la primera mitad del siglo XVIII, por G. Monge (1785), y en la primera parte del siglo XIX por A. L. Cauchy (1826). Sin embargo, la primera exposición *sistemática* de una geometría multidimensional (más de tres dimensiones) fue dada en 1844 por el matemático alemán H. Grassmann e independientemente por los ingleses W. R. Hamilton y A. Cayley apoyándose en una analogía formal con la geometría analítica ordinaria.

a. Rectas en $\mathbb{R}^n$

Observemos que el conjunto de puntos X que están sobre una línea recta en $\mathbb{R}^n$ que pasa por un punto P y que es *paralela* a un vector $N \neq 0$ satisfacen la condición de que $X - P$ es paralelo a N; es decir, existe $t \in \mathbb{R}$ tal que $X - P = tN$; o, en otra forma,

$$X = P + t\,N \tag{1}$$

donde $t \in \mathbb{R}$ (figura 27). Esta igualdad se denomina *ecuación paramétrica* de la recta en $\mathbb{R}^n$, y al vector N se le llama *vector director* de la recta.

Observemos que si $n = 2$, la recta está en el plano cartesiano; y si $n = 3$, la recta está en el espacio tridimensional. Si el lector no tiene algún problema en considerar espacios de más de tres dimensiones, podrá observar que puede estudiar también ecuaciones de rectas en espacios de dimensiones $n = 4, 5, \ldots$, utilizando la ecuación (1).

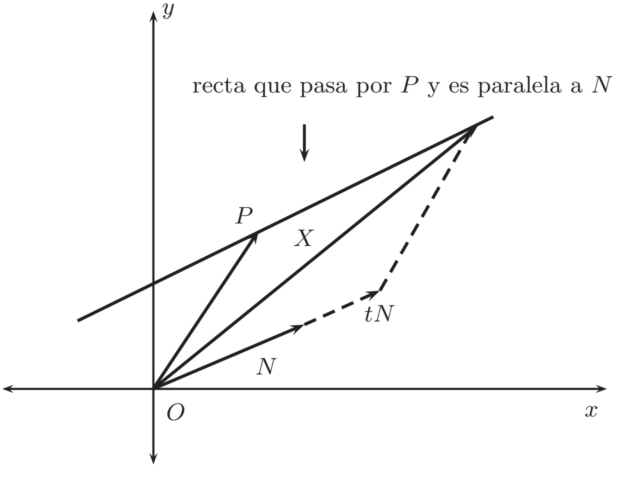


Figura 27. Ecuación paramétrica de una recta

Ejemplo 17. (Ecuación de una recta en $\mathbb{R}^2$)
Consideremos el plano $\mathbb{R}^2$ y representemos un punto X sobre la recta L por sus coordenadas cartesianas (x, y). Sean $P = (p, q)$ y $N = (a, b)$. Entonces, en términos de las coordenadas x y y, la ecuación (1) de la recta $X = P + tN$ se puede expresar mediante las ecuaciones

$$x = p + ta$$

$$y = q + tb$$

donde $t \in \mathbb{R}$. Si $a \neq 0$ y $b \neq 0$, podemos, todavía más, eliminar el parámetro t de estas dos igualdades y obtener *la ecuación cartesiana* de la recta:

$$\frac{x-p}{a} = \frac{y-q}{b} \qquad \text{ó} \qquad y = \frac{b}{a}x + \left[q - \frac{b}{a}p\right]$$

de manera que la pendiente estaría dada por $m = \dfrac{b}{a}$; es decir, por el cociente de las componentes del vector normal.

Ejemplo 18.
Sean $P = (3, 1)$ y $N = (-1, 2)$. Entonces una ecuación paramétrica de la recta en $\mathbb{R}^2$ que pasa por P en la dirección de N es, para $t \in \mathbb{R}$,

$$x = 3 - t, \qquad y = 1 + 2t$$

La ecuación cartesiana podemos expresarla entonces como

$$y = -2x + 7$$

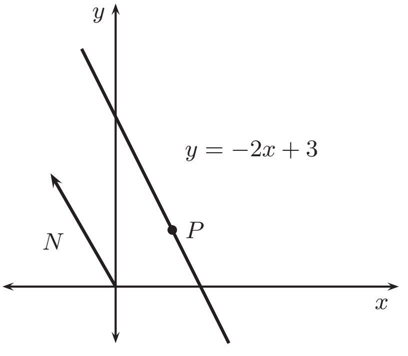


Figura 28

Nota 9. (Más fácil con teoría de vectores)
Si $y = m_1x + b_1$ y $y = m_2x + b_2$ son dos rectas en el plano, entonces sus vectores paralelos pueden ser, respectivamente, $(1, m_1)$ y $(1, m_2)$. Luego la condición para que las rectas sean ortogonales (o perpendiculares) es que $(1, m_1) \cdot (1, m_2) = 0$ o, equivalentemente, $m_1m_2 = -1$ (Volumen 0 (Fundamentos)).

Ejemplo 19. (Ecuación de una recta en $\mathbb{R}^3$)
Consideremos el plano $\mathbb{R}^3$ y representemos un punto X sobre la recta L por sus coordenadas cartesianas (x, y, z). Sean $P = (p, q, r)$ y $N = (a, b, c)$. Entonces, en términos de las coordenadas x, y y z, la ecuación de la recta $X = P + tN$ se puede expresar mediante las ecuaciones

$$\begin{aligned} x &= p + ta \\ y &= q + tb \\ z &= r + tc \end{aligned}$$

donde $t \in \mathbb{R}$. Si $a \neq 0$, $b \neq 0$ y $c \neq 0$ podemos eliminar el parámetro t de estas tres igualdades y obtener *la ecuación cartesiana* de la recta:

$$\frac{x - p}{a} = \frac{y - q}{b} = \frac{z - r}{c}$$

Ejemplo 20.
Determinemos la recta en $\mathbb{R}^3$ que pasa por los puntos $A = (1, 2, -1)$ y $B = (-1, -1, 1)$.

Solución

Un vector que señala la dirección de esta recta lo podemos calcular como $N = A - B = (1, 2, -1) - (-1, -1, 1) = (2, 3, -2)$. Una ecuación paramétrica de la recta es (con $P = A$)

$$X = P + tN = (1, 2, -1) + t(2, 3, -2) \tag{2}$$

donde $t \in \mathbb{R}$. Para encontrar una ecuación en términos de las coordenadas cartesianas, hagamos $X = (x, y, z)$. Entonces las ecuaciones paramétricas son:

$$\begin{aligned} x &= 1 + 2t \\ y &= 2 + 3t \qquad t \in \mathbb{R} \\ z &= -1 - 2t \end{aligned}$$

Despejando t e igualando, se tiene que la ecuación cartesiana de la recta es

$$\frac{x-1}{2} = \frac{y-2}{3} = \frac{-z-1}{2}$$

Un ejercicio aquí sería que el lector comprobara que se obtiene la *misma* recta si en lugar de $P = A$, hubiéramos escogido $P = B$, en la ecuación (2) arriba.
▲

Ejemplo 21.
Probemos que la ecuación paramétrica de la recta en $\mathbb{R}^3$ que pasa por los puntos $(1, -1, 2)$ y $(7, 0, 5)$ es, para $t \in \mathbb{R}$,

$$\begin{aligned} x &= 6t + 1 \\ y &= t - 1 \qquad t \in \mathbb{R} \\ z &= 3t + 2 \end{aligned}$$

Solución

Un vector que da la dirección de esta recta lo podemos calcular como $N = (7, 0, 5) - (1, -1, 2) = (6, 1, 3)$. Una ecuación paramétrica de la recta es, entonces,

$$X = (1, -1, 2) + t(6, 1, 3)$$

donde $t \in \mathbb{R}$. Para encontrar una ecuación en términos de las coordenadas cartesianas, hagamos $X = (x, y, z)$. Entonces

$$\begin{aligned} x &= 1 + 6t \\ y &= -1 + t \qquad t \in \mathbb{R} \\ z &= 2 + 3t \end{aligned}$$

Despejando el parámetro t e igualando se tiene que

$$\frac{x-1}{6} = y + 1 = \frac{z-2}{3}$$

b. Planos en $\mathbb{R}^n$

Sea $P \in \mathbb{R}^n$ un punto, y tomemos un vector $N \in \mathbb{R}^n, N \neq 0$. El *plano* que pasa por el punto P y es perpendicular al vector N puede definirse como el conjunto de puntos $X \in \mathbb{R}^n$, tales que

$$(X - P) \cdot N = 0 \qquad \text{(ecuación del plano en } \mathbb{R}^n) \qquad (1)$$

Notemos que en la ecuación anterior se exige que el vector $X - P$ sea perpendicular al vector N (figura 29).

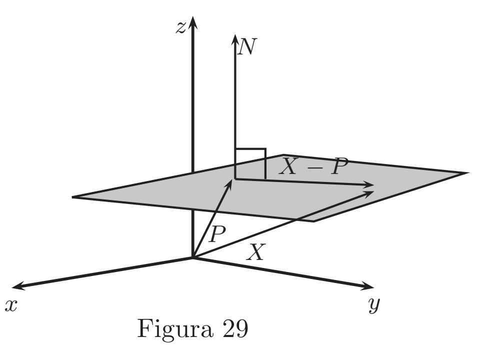


Figura 29

Ejemplo 22. (Ecuación de un plano en $\mathbb{R}^3$)
Para escribir la ecuación cartesiana de un plano en $\mathbb{R}^3$, tomemos los vectores involucrados en términos de sus componentes; es decir, $X = (x, y, z)$, $N = (a, b, c)$ y $P = (x_0, y_0, z_0)$. Entonces la ecuación del plano $(X - P) \cdot N = 0$ es

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

Ejemplo 23.
La ecuación cartesiana del plano en $\mathbb{R}^3$ que pasa por el punto $P = (2, -1, 3)$ y es perpendicular al vector $N = (1, 1, 1)$ es $1(x-2) + 1(y+1) + 1(z-3) = 0$ o, lo que es equivalente,

$$x + y + z = 4$$

Ejemplo 24. (Plano que pasa por tres puntos)
Encontremos la ecuación del plano en $\mathbb{R}^3$ que pasa por los puntos $P_1 = (1, -1, 2)$, $P_2 = (3, 1, 0)$, $P_3 = (-1, 2, 7)$.

Solución
Tenemos $N = \overrightarrow{P_3P_1} \times \overrightarrow{P_2P_1}$, donde $\overrightarrow{P_3P_1} = P_1 - P_3$ y $\overrightarrow{P_2P_1} = P_1 - P_2$. Entonces,

$$\overrightarrow{P_3P_1} = (2, -3, -5)$$
$$\overrightarrow{P_2P_1} = (-2, -2, 2)$$
$$N = \overrightarrow{P_3P_1} \times \overrightarrow{P_2P_1} = \begin{vmatrix} i & j & k \\ 2 & -3 & -5 \\ -2 & -2 & 2 \end{vmatrix} = (-16, 6, -10)$$

Por tanto, la ecuación del plano que pasa por los puntos $P_1 = (1, -1, 2)$, $P_2 = (3, 1, 0)$ y $P_3 = (-1, 2, 7)$ es (tomando a P_1 como el punto por el cual pasa el plano)

$$-16(x-1) + 6(y+1) - 10(z-2) = 0$$

que es equivalente a

$$8x - 3y + 5z = 21 \quad \blacktriangle$$

Podemos ahora resumir algunas características geométricas de las rectas y los planos, utilizando los conceptos previamente presentados. Veamos esto.

Definición 14. (Sobre rectas y planos)

a) Dos *líneas rectas* en $\mathbb{R}^n$ se dicen *paralelas* si dados dos puntos distintos P_1, Q_1 sobre la primera línea, y otros dos puntos P_2, Q_2 sobre la segunda, los vectores $P_1 - Q_1$ y $P_2 - Q_2$ son paralelos (figura 30).

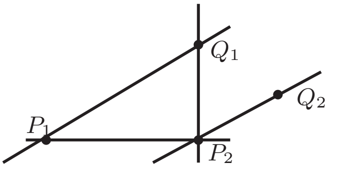


Figura 30. Rectas paralelas

b) Dos *planos* en $\mathbb{R}^n$ son *paralelos* si sus vectores normales son paralelos (figura 31).

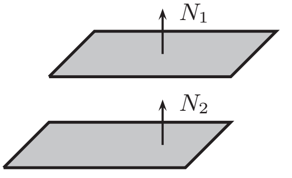


Figura 31. Planos paralelos

c) Dos *planos* en $\mathbb{R}^n$ son *perpendiculares* si sus vectores normales son perpendiculares (figura 32).

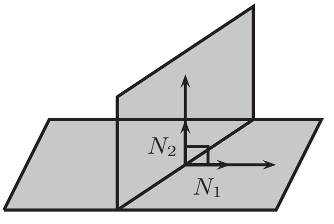


Figura 32. Planos perpendiculares

d) El *ángulo entre dos planos* en $\mathbb{R}^n$ se define como el ángulo entre sus vectores normales (figura 33).

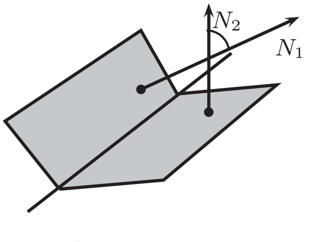


Figura 33. Ángulo entre planos

Ejemplo 25.

a) El plano $3x + 5y + 2z = 1$ es paralelo al plano $6x + 10y + 4z = -1$, pues $(3, 5, 2) = \frac{1}{2}(6, 10, 4)$.

b) El plano $7x - 5y + z = 3$ es perpendicular al plano $5x + 6y - 5z = 8$, pues $(7, -5, 1)(5, 6, -5) = 0$.

c) El plano $x-2y+3z=2$ es paralelo al plano $3x-6y+9z=-2$.

d) El plano $4x+2y+2z=5$ es perpendicular al plano $x-3y+z=3$.

Ejercicios 4

1) Determine una ecuación paramétrica para la recta que pasa por el punto P y es paralela al vector N, donde

   a) $P=(1,2)$ y $N=(3,4)$.
   b) $P=(0,0)$ y $N=(5,7)$.
   c) $P=(3,1,2)$ y $N=(7,-1,-2)$.

   En cada caso, obtenga la ecuación cartesiana.

2) Encuentre las ecuaciones paramétricas de la línea que satisface las condiciones dadas:

   a) Pasa a través de $(2,1,1)$ y $(3,2,6)$.
   b) Pasa a través de $(1,1,3)$ y es paralela a la línea $x=2-6t$, $y=1+4t$, $z=7t$.
   c) Pasa a través de $(0,0,0)$ y es perpendicular al plano $6x-2y+z=3$.
   d) Pasa a través de $(1,2,2)$ y es perpendicular a la recta $x=2+t$, $y=1-t$, $z=5+t$ y a la recta $x=2+t$, $y=3+2t$, $z=9+3t$.

3) Encuentre la ecuación del plano que satisface las condiciones dadas:

   a) Pasa a través de los puntos $(0,0,1)$, $(1,3,7)$ y $(-1,2,3)$.
   b) Pasa a través de $(1,2,2)$ y es perpendicular a la línea $x=2-t$, $y=3+2t$, $z=2-8t$.
   c) Pasa a través de la línea $x=1+t$, $y=2+3t$, $z=1+6t$ y es perpendicular al plano $3x+2y-5z=0$.

4) Muestre que las ecuaciones paramétrica y cartesiana de la recta que pasa por $(2,-9,5)$ y es paralela a $(0,2,3)$ son

$$x=2, \qquad y=-9+2t, \qquad z=5+3t, \qquad t\in\mathbb{R}$$

   y

$$x=2, \qquad \frac{y+9}{2}=\frac{z-5}{3}$$

5) Muestre que la ecuación cartesiana del plano que pasa por $P=(-3,0,7)$ y es perpendicular al vector $N=(5,2,-1)$ es $z-2y-5x=22$.

6) Pruebe que la ecuación del plano que contiene a la línea $x - 2y + z = 1$, $2x - y + z = 0$ y es perpendicular al plano $3x + 2y - 3z = 0$, está dada por $9x - 6y + 5z = 1$.

7) En la lección 3 (*"Geometría analítica"*) del volumen 0 (Fundamentos), afirmábamos que las siguientes son ecuaciones paramétricas de la recta que pasa por los puntos $A(x_0, y_0)$ y $B(x_1, y_1)$:

$$x = x_0 + \frac{t(x_1 - x_0)}{\sqrt{(x_0 - x_1)^2 + (y_0 - y_1)^2}}$$

$$y = y_0 + \frac{t(y_1 - y_0)}{\sqrt{(x_0 - x_1)^2 + (y_0 - y_1)^2}}$$

Escriba éstas en la forma vectorial estándar $X = P + tN$ que estudiamos en esta lección. ¿Qué relación existe entre P, N y A, B?

5. Contexto económico

a. El modelo de equilibrio general Walras-Cassel (1918)

El *modelo de equilibrio general Walras-Cassel* (1874, 1918), como una particular abstracción de los procesos de producción de una economía, tiene una importante tradición en la historia de la teoría económica desde (quizás) Quesnay (y su *Tableau Économique* de 1758), pasando por Ricardo, Marx y el mismo Walras, hasta nuestros días.

Léon Walras, en la primera y segunda ediciones de *Élements d'Economie Politique Pure* (1874, 1877) tenía en mente una economía en la que cada mercancía tiene un conjunto de procesos (o actividades) de producción; cada proceso convierte una unidad de cada insumo en una cantidad dada de otra mercancía; y luego esta actividad se expande o contrae a través de "rendimientos constantes a escala" (con coeficientes fijos de fabricación), es decir, linealmente (si una unidad de cada insumo produce x unidades de mercancía, entonces dos unidades producirán $2x$ unidades de mercancía; y $\frac{1}{2}$ unidad de cada insumo producirá $\frac{1}{2}x$ unidades de mercancía, etc.). Ninguna actividad interactúa con otra, excepto por el uso del recurso común.

Una característica del pensamiento de Walras es que pareciera operar como si la producción transformara recursos primarios en bienes finales. Pero, de hecho, los bienes intermedios (o de capital) los acomoda a través de un esquema de solución de cierto conjunto de ecuaciones lineales simultáneas. De todas maneras, aún cuando Walras trabajó con coeficientes fijos, tenía en mente un esquema mucho más amplio. En la segunda edición de sus *Elements* (1877), por ejemplo, aseguraba que los coeficientes fijos eran sólo una conveniencia; que la sustitución técnica era posible; y que los coeficientes dependían de los precios de los insumos. Y en la tercera edición ya utilizaba funciones de producción, y coeficientes de fabricación variables (determinados por un proceso de minimizar costos) como incógnitas de un sistema de *equilibrio general.*

Pero Walras seguía hablando de los coeficientes de fabricación (insumos por unidad de producto) variables como cambios en las proporciones de insumos, pero constantes para cambios en producción: *Walras seguía asumiendo rendimientos constantes a escala con coeficientes constantes, y sus métodos, parcialmente, deberían haber sido los de una economía lineal, aunque esto nunca ocurrió así, quizás debido a su débil formación matemática.*

Gustave Cassel (1918)[1], quien fuera el primero en *popularizar* el sistema walrasiano, que había caído en desuso desde entonces, utilizó el modelo de los coeficientes fijos e introdujo otras simplificaciones. Planteaba entonces el primer

[1] Cassel, Gustav (1918), *Theory of Social Economy*, edición de 1932; New York: Harcourt, Brace and Company, 1932 edn.

problema básico del sistema walrasiano abstracto: un sistema de ecuaciones de oferta igualadas a ecuaciones de demanda en cada uno de los factores debía ser resuelto por un sistema de precios (uno para cada factor). Este es el problema de la *existencia de una solución de equilibrio.* El segundo problema es el de la *unicidad del equilibrio* (si existe). Y agregándole cierta dinámica, el tercer problema es describir la respuesta de los precios (y sus cantidades asociadas) a situaciones de desequilibrio: este es el problema de la *estabilidad del equilibrio.*

Walras sabía que estas eran preguntas legítimas e importantes y creía haber hecho mucho por resolverlas. Sin embargo, lo suyo sobre esto eran simples notas que estaban lejos de una solución rigurosa. Por ejemplo, menospreciaba el problema de la existencia del equilibrio y su unicidad diciendo que su sistema contenía exactamente el mismo número de ecuaciones que de incógnitas[2], y esto, se sabe, no es ni condición suficiente ni tampoco necesaria para la unicidad de soluciones: el conteo de ecuaciones no aclara nada.

Cassel considera una economía con n mercancías (bienes finales) y m factores (insumos) de producción. Sea r_i la cantidad ofrecida del factor i y sea x_j la cantidad producida de la mercancía j. Las posibilidades técnicas de la producción están caracterizadas por mn coeficientes fijos a_{ij}, que representan la cantidad física del factor i-ésimo utilizado en la fabricación de una unidad de la mercancía j-ésima.[3] De esta forma, la demanda total del factor i-ésimo es entonces $a_{i1}\,x_1 + a_{i2}\,x_2 + \cdots + a_{in}\,x_n$. Ahora: si igualamos la oferta a la demanda en cada uno de los factores, obtenemos m ecuaciones de equilibrio

$$\begin{array}{ccccccc}
a_{11}\,x_1 & + & a_{12}\,x_2 & + \cdots + & a_{1n}\,x_n & = & r_1 \\
a_{21}\,x_1 & + & a_{22}\,x_2 & + \cdots + & a_{2n}\,x_n & = & r_2 \\
\vdots & & \vdots & & \vdots & & \vdots \\
a_{m1}\,x_1 & + & a_{m2}\,x_2 & + \cdots + & a_{mn}\,x_n & = & r_m
\end{array}$$

o, en forma matricial,

$$Ax = r \tag{1}$$

donde $A = (\,a_{ij}\,)_{m\times n},\ x = [\,x_1, \ldots, x_n\,]^T,\ r = (\,r_1, \ldots, r_m\,)^T$ [4].

De otro lado, sean $p_1, \ldots, p_n$ los precios de las mercancías, y sean $v_1, \ldots, v_m$ los precios de los m factores. Las ecuaciones de demanda del mercado por las

---

[2] En este punto es conveniente aclarar que aunque los procesos presenten coeficientes fijos de fabricación, las ofertas y las demandas de los distintos factores no son, necesariamente, funciones lineales, como veremos enseguida.

[3] Aunque el modelo Walras-Cassel analiza una economía en la cual ciertos "bienes primarios" se transforman en bienes finales, es posible extenderlo para incluir bienes intermedios.

[4] Observemos que si $m > n$ (es decir, si el número de factores es mayor que el número de productos) este sistema podría no tener solución.

mercancías deben poderse escribir mediante las condiciones de equilibrio:

$$\begin{aligned} x_1 &= d_1(\,p_1,\ldots,p_n;v_1,\ldots,v_m\,) \\ x_2 &= d_2(\,p_1,\ldots,p_n;v_1,\ldots,v_m\,) \\ \vdots & \qquad\quad \vdots \\ x_n &= d_n(\,p_1,\ldots,p_n;v_1,\ldots,v_m\,) \end{aligned} \tag{2}$$

donde $d_1(\cdot),\ldots,d_n(\cdot)$ son *funciones de demanda homogéneas de grado cero*, es decir, para todo $t>0$,

$$d_i(\,tp_1,\ldots,tp_n;tv_1,\ldots,tv_m\,) = d_i(\,p_1,\ldots,p_n;v_1,\ldots,v_m\,)$$

para $i=1,2,\ldots,n$ [5].

Ahora: la relación de equilibrio (que el mismo Walras llamara de "competencia perfecta") entre precios de factores y precios de bienes finales (o mercancías) es *directa* pues no existen, en este modelo, bienes intermedios (o de capital):

$$\begin{aligned} a_{11}\,v_1 + a_{21}\,v_2 + \cdots + a_{m1}\,v_m &= p_1 \\ a_{12}\,v_1 + a_{22}\,v_2 + \cdots + a_{m2}\,v_m &= p_2 \\ \vdots \qquad \vdots \qquad\quad \vdots \qquad & \quad \vdots \\ a_{1n}\,v_1 + a_{2n}\,v_2 + \cdots + a_{mn}\,v_m &= p_n \end{aligned}$$

o, en forma matricial,

$$A^T v = p \tag{3}$$

donde $v = (\,v_1,v_2,\ldots,v_m\,)^T,\ p = (\,p_1,p_2,\ldots,p_n\,)^T$.

Finalmente, lo que se necesita para redondear el sistema Walras-Cassel es cierta consideración con respecto a la oferta de recursos; es decir, que la oferta de recursos dependa de los precios de los bienes finales y de los precios de los factores:

$$\begin{aligned} r_1 &= g_1(\,p_1,\ldots,p_n;v_1,\ldots,v_m\,) \\ r_2 &= g_2(\,p_1,\ldots,p_n;v_1,\ldots,v_m\,) \\ \vdots & \qquad\quad \vdots \\ r_m &= g_m(\,p_1,\ldots,p_n;v_1,\ldots,v_m\,) \end{aligned} \tag{4}$$

[5] Obsérvese que Cassel (a diferencia de Walras) no recurrió a las *funciones de utilidad* sino que dirigió directamente su atención a las *demandas*.

donde las funciones $g_i(\cdot)$ son homogéneas de grado cero en sus argumentos; es decir, para todo $t > 0$,

$$g_i(\, tp_1, \ldots, tp_n; tv_1, \ldots, tv_m\,) = g_i(\, p_1, \ldots, p_n; v_1, \ldots, v_m\,)$$

De estos cuatro sistemas de ecuaciones se encuentra una relación que ha venido a jugar un papel fundamental en la teoría del equilibrio general: *la ley de Walras*[6]. Observemos que si multiplicamos a ambos lados de las ecuaciones del sistema (1) por $[\, v_1, \ldots, v_m\,]^T$, y a ambos lados de las ecuaciones del sistema (3) por $[\, x_1, \ldots, x_n\,]^T$, obtenemos que

$$p \cdot x = \sum_{i=1}^{n} p_i x_i = \sum_{j=1}^{m} v_j r_j = v \cdot r \qquad \text{(Ley de Walras)}$$

Esta *ecuación de equilibrio* (que no es más que cierto tipo de “restricción presupuestal”) afirma que, en el agregado, la valoración de la demanda iguala a la valoración de la oferta en términos de la unidad monetaria de los precios. Observemos que de la ley de Walras se deduce que, en realidad, *sólo* $m+n-1$ *ecuaciones del modelo son las fundamentales*: si $m+n-1$ ecuaciones de oferta-demanda se satisfacen, entonces el total de $m + n$ ecuaciones igualmente se satisfarán.

Otra observación es fundamental: si multiplicamos los precios de los factores (v_i’s) y los precios de los productos finales (p_i’s) por un escalar $t > 0$, los sistemas (1), (2), (3) y (4) se mantienen idénticos. La manera walrasiana de manejar esto es elegir una mercancía, digamos la primera, como *numerario* ($p_1 = 1$) y reducir así el número de incógnitas en uno. El sistema se resuelve entonces para precios relativos al precio del numerario: *las economías walrasianas no pueden encontrar precios absolutos, solo precios relativos al precio numerario escogido arbitrariamente.* Es el numerario elegido arbitrariamente el que “cierra” el modelo.

Una observación más: notemos que en el sistema Walras-Cassel, ni los precios determinan los costos de producción, ni éstos determinan los precios. En equilibrio, el precio iguala al costo de producción, y éstos se obtienen como soluciones a un *sistema simultáneo* y no por dirección causal como afirmaran los clásicos.

Está claro que el sistema (no necesariamente lineal) de Walras-Cassel que acabamos de presentar muestra que *no* se puede concluir (como sí lo hiciera

[6] Llamada así por Óscar Lange en su artículo *“Say´s Law: A Restatement and Criticism”* de 1942, aparecido en Lange et ál., (eds.), *Studies in Mathematical Economics.*

Walras) que puesto que tenemos $m+n$ ecuaciones con $m+n$ incógnitas, entonces la solución tiene exactamente una solución y que esta solución tiene significado económico (todos los precios y cantidades son no-negativos). El primer estudio riguroso del sistema Walras-Cassel lo hizo Abraham Wald en una secuencia de artículos de los años treinta del siglo pasado (publicados resumidos y sin pruebas en *Econometrica* de octubre de 1951 bajo el título *Some Systems of Equations of Mathematical Economics*). Allí investigaba la existencia y unicidad de soluciones del sistema Walras-Cassel bajo algunas hipótesis no del todo satisfactorias que disminuyeron la validez de sus aportes. Aún así, fue un gran logro: fue quizás el trabajo más completo y riguroso en economía matemática hasta esa fecha, puesto que todavía el modelo de crecimiento de von Neumann de 1932 no se conocía ampliamente.

La *apariencia lineal* del sistema Walras-Cassel (los sistemas (1) y (3) son lineales) fue lo que llevó a pensar a los pioneros en una solución obvia. Pero los sistemas de ecuaciones (2) y (4) (características no necesariamente lineales de las funciones de demanda y oferta) aclaran que este problema no puede ser atacado, con toda generalidad, con técnicas básicas del álgebra lineal: serán necesarias herramientas del análisis matemático y de la topología[7]. Éste fue el trabajo de Kenneth Arrow [1921-] (Premio Nobel de economía en 1972), Gerard Debreu [1921-2004] (Premio Nobel de economía en 1983), y Tjalling Koopmans [1910-1985] (Premio Nobel de economía en 1975), entre otros, a mediados del siglo XX. Sobre este último discutiremos en la lección 8.

Ejemplo 26. (Economía en equilibrio según Cassel y Wald)
Supongamos que $a_{11}=1$, $a_{12}=2$, $a_{21}=4$, $a_{22}=5$, $r_1=10$, $r_2=30$ (así, hay dos bienes finales y dos recursos). Supongamos, además, que las demandas están dadas por $x_1=\dfrac{3}{p_1}$, y $x_2=\dfrac{4}{p_2}$. Entonces los sistemas (1), (2) y (3) se reducen a

$$\begin{aligned} x_1+2x_2&=10\\ 4x_1+5x_2&=30 \qquad (1)\\ x_1=\frac{3}{p_1} \qquad x_2&=\frac{4}{p_2} \qquad (2)\\ v_1+4v_2&=p_1\\ 2v_1+5v_2&=p_2 \qquad (3)\end{aligned}$$

De (1) observemos que $x_1=x_2=\frac{10}{3}$; y así, de (2), $p_1=\frac{9}{10}$, $p_2=\frac{6}{5}$, y estas últimas, insertadas en (3), arrojan que $v_1=\frac{1}{10}$, $v_2=\frac{1}{5}$. Estas son, en este

[7] Ver Volumen 2 (Cálculo) y Volumen 3 (Optimización y dinámica).

caso, las soluciones al sistema de equilibrio general Walras-Cassel (el lector podría preguntarse ahora por qué el precio del factor 1 es la mitad del precio del factor 2).

Obsérvese de nuevo cómo, aquí, *no* es el costo de producción el que determina los precios como afirmaran los clásicos, sino que son los precios de los productos los que determinan el costo de producción. Esto está, además, en consonancia con el *principio de imputación* de von Wieser (1889) que analizábamos en la lección 3 anterior.▲

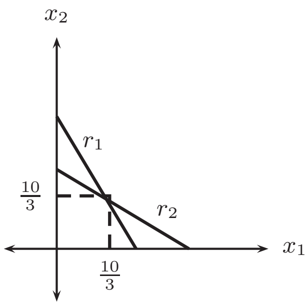


Figura 34

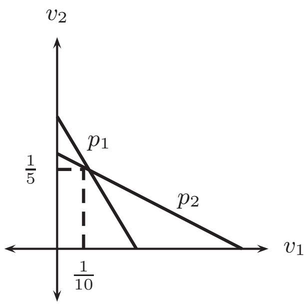


Figura 35

Ejemplo 27. (Un bien "escaso" con precio nulo)
Supongamos una economía con coeficientes técnicos $a_{11} = 1$, $a_{12} = 2$, $a_{21} = 4$, $a_{22} = 5$, y recursos $r_1 = 87/11$, $r_2 = 30$. Supongamos, además, que las demandas están dadas por $x_1 = \dfrac{10}{p_1}$, y $x_2 = \dfrac{1}{p_2}$. Entonces los sistemas (1), (2) y (3) se reducen a

$$\begin{aligned} x_1 + 2x_2 &= 87/11 \\ 4x_1 + 5x_2 &= 30 \end{aligned} \tag{1}$$

$$x_1 = \frac{10}{p_1} \quad x_2 = \frac{1}{p_2} \tag{2}$$

$$\begin{aligned} v_1 + 4v_2 &= p_1 \\ 2v_1 + 5v_2 &= p_2 \end{aligned} \tag{3}$$

Resolviendo el sistema (1) obtenemos que $x_1 = \frac{75}{11}, x_2 = \frac{6}{11}$; y así, de (2), $p_1 = \frac{11}{30}$, $p_2 = \frac{11}{6}$; y estas últimas, insertadas en (3), arrojan que $v_1 = 0$, $v_2 = \frac{22}{15}$. ¿Podría el lector explicar (económicamente) por qué ocurre esto aquí? Es decir, ¿por qué $v_1 = 0$ si el recurso r_1 es "escaso"?

Nota 10. (Un llamado final)
Algunos economistas creen que la visión del modelo Walras-Cassel no se adhiere a la visión original del propio Walras, ni en propósito ni en detalle, y llaman *neo-walrasiana* a esta vertiente que surge del modelo de Cassel. Una de las razones para esto es que se ha ignorado, de la obra de Walras, piezas que él mismo consideraba fundamentales. A *Studies in Social Economics* (1896) y *Studies in Applied Economics* (1898) las consideraba, no como meras compilaciones de trabajos previos, sino como libros complementarios de los *Élements*. De hecho, los *Élements* de 1874 los subtituló *Theory of Social Wealth*; mientras que su libro de 1896 lo subtituló *Theory of the Division of Social Wealth*; y el de 1898, *Theory of the Production of Social Wealth*. La visión que hoy llamamos *walrasiana* sólo abarca (por razones que únicamente la historia podrá evaluar) sus *Élements* de 1874. Sobre esto regresaremos en la lección 2 del volumen 3: Optimización y dinámica.

Ejercicios complementarios

1) Si $x = (2, 1, 5)$, $y = (3, -1, 2)$, $z = (-5, 1, 1)$, encuentre

   a) $x \cdot y, \ \ y \cdot z$

   b) $||\, x\,||, \, ||\, y\,||, \, ||\, z\,||$

   c) $(x + y) - z$

   d) $||\, x + y + z\,||, \, ||\, x + y\,||, \ ||\, z\,||$

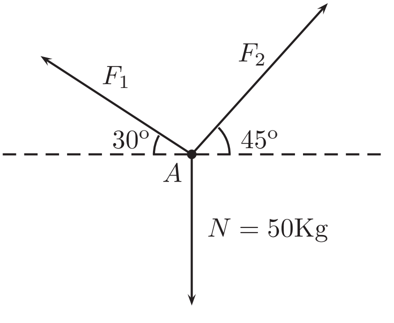
2) Si $x = (-1, -1, -1)$, $y = (2, 1, 3)$, $z = (4, 0, 4)$, encuentre el ángulo entre los siguientes pares de vectores:

   a) $x, \, y$ b) $x, \ y + z$

   c) $x + y, \ z$ d) $x + y, \ x - y$

   e) $x \times y, \, z$ f) $y \times z, \, x$

3) Considere el siguiente sistema mecánico: tres fuerzas actúan en el punto A de la figura de abajo. Calcule las fuerzas F_1 y F_2 de tal forma que el punto A esté en equilibrio; es decir, que $F_1 + F_2 = N$.



   [Indicación: escriba F_1 y F_2 como parejas ordenadas, y resuelva el sistema de ecuaciones lineales resultante.]

4) Encuentre el ángulo entre las rectas $4x - y = 5$ y $7x + 2y = 1$. [Indicación: calcule vectores paralelos N a cada una de ellas].

5) Encuentre el ángulo entre los planos $x + 3y - z = 1$ y $2x + 7y + 5z = 8$.

6) Encuentre todos los vectores que son perpendiculares a $(1, 3, 3, 1)$ y $(2, 9, 8, 1)$.

7) Si $x \in \mathbb{R}^n$ y $x \cdot y = 0$ para todo y, pruebe que $x = 0$.

8) Encuentre los ángulos del triángulo en $\mathbb{R}^3$ con vértices $(2,1,6)$, $(8,1,4)$, $(3,2,9)$.

9) Muestre que la recta que pasa por $(1,3)$ y es perpendicular a la recta $x-2y+2=0$ es $y=-2x+5$.

10) Encuentre el plano en el espacio que pasa por los tres puntos dados:

    a) $(1,6,1),(9,1,-32),(-4,-2,50)$
    b) $(4,0,0),(0,4,0),(0,0,4)$
    c) $(2,1,1)$, $(3,-1,1)$ y $(4,1,2)$

11) Encuentre un vector perpendicular al plano $4x+5y+9z=1$.

12) Demuestre que las rectas $3x+5y=1$ y $15x-9y=7$ son ortogonales.

13) ¿Para qué valores de c son ortogonales los planos $x+2y+3z=6$, $cx+y+z=9$?

14) Sólo utilizando vectores, demuestre que si las diagonales de un rectángulo son ortogonales, el rectángulo debe ser un cuadrado.

15) ¿Está $(2,8,1)$ en la recta que pasa por $(1,2,3)$ y $(9,2,0)$?

16) Encuentre el punto medio del segmento de recta formado por $x=(2,0,1)$ y $y=(10,4,3)$. (Indicación: encuentre z tal que $\|\overrightarrow{xz}\| = \|\overrightarrow{yz}\|$).

17) Pruebe que $x-y$ es ortogonal a $x+y$ si, y sólo si, $\|x\|=\|y\|$. ¿Cuál es el significado geométrico de esto?

18) Pruebe que el plano $x+2y+z=1$ es ortogonal a la recta que pasa por el origen a través del punto $(1,2,1)$.

19) a) ¿Cuál es la intersección de los planos $x-y+3z=5$ y $2x+2y+7z=8$ en $\mathbb{R}^3$?

    b) ¿Y cuál es la intersección de los planos $x+y+5z=7$ y $2x+2y+10z=9$?

20) ¿Cuál es la intersección de los hiperplanos $x+y+z+w=0$ y $x+y-z-w=0$ en $\mathbb{R}^4$?

21) ¿Será que los vectores $(1,2,1)$, $(0,1,3)$, $(5,1,8)$ y $(-1,0,2)$ están en el mismo plano de $\mathbb{R}^3$?

22) Muestre que la condición para que la línea

$$\frac{x-a}{l}=\frac{y-b}{m}=\frac{z-c}{n}$$

sea paralela al plano $Ax+By+Cz+D=0$ es que $Al+Bm+Cn=0$. ¿Cuál es la condición para que la recta esté *sobre* el plano?

23) ¿Están los vectores $(7,2,9)$, $(0,1,3)$ y $(7,7,6)$ en el mismo plano en $\mathbb{R}^3$? Si es así, encuentre un vector ortogonal a ese plano.

24) Encuentre un vector paralelo a la línea de intersección de los planos:

a) $2x-y+z=1,\ 3x+y+z=2$

b) $3x-5y+z=2,\ x-2y+2z=4$

25) Encuentre el ángulo entre los siguientes planos:

a) $\begin{aligned} x+y+z&=2\\ -x-y+z&=7\end{aligned}$ b) $\begin{aligned} 5x-y-z&=9\\ 10x+y+8z&=1\end{aligned}$

26) Pruebe que la ecuación del plano que pasa por el punto $(1,-2,3)$ y es perpendicular a la línea de intersección de los planos $3x+2y-2z=12$ y $x+2y+2z=0$ es $2x-2y+z=9$.

27) Pruebe que la ecuación del plano que contiene los puntos $(1,-3,2)$, $(2,-4,-5)$ y $(3,-2,0)$ es $3x-4y+z=17$.

28) Pruebe que la ecuación del plano perpendicular al segmento de línea que va desde $(2,-3,4)$ hasta $(6,3,2)$ y que pasa por el punto medio de este segmento es $2x+3y-z=5$.

29) Muestre que la ecuación cartesiana de la línea $3x+2y-2z=10,\ 3x-4y+z=1$ se puede escribir como

$$\frac{x-2}{2}=\frac{y-1}{3}=\frac{z+1}{6}$$

¿Cuál es su forma paramétrica?

30) Pruebe que la línea que pasa por $(3,-2,4)$ y $(5,-5,3)$ está dada por

$$\frac{x-3}{2}=-\frac{y+2}{3}=-\frac{z-4}{1}$$

31) Muestre que la línea

$$\frac{x+1}{2} = -\frac{y-3}{1} = \frac{z+2}{3}$$

es paralela al plano $3x + 3y - z + k = 0$. Encuentre k tal que el plano *contenga* la línea.

32) Muestre que la línea cuyas ecuaciones paramétricas son

$$x = 1 - \frac{2}{15}t, \qquad y = 3 + \frac{1}{3}t, \qquad z = 2 - \frac{14}{15}t$$

es perpendicular al plano $2x - 5y + 14z = 0$.

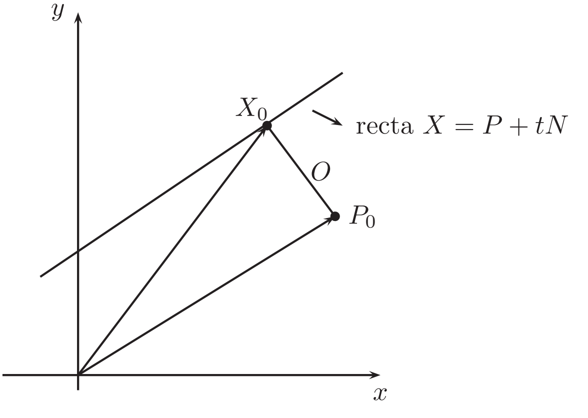
33) Pruebe que el vector (x_0, y_0, z_0) está en el plano generado por $(4, 3, -1)$ y $(3, -2, 12)$ si, y sólo si, $z_0 = 2x_0 - 3y_0$.

34) *(Distancia a rectas y planos)*

a) Tomemos una recta en el plano (o en el espacio) de la forma estándar $X = P + tN$, y un punto P_0 (en el plano o en el espacio, respectivamente). Observe (ver figura) que la distancia del punto P_0 a la recta estará dada por

$$d = ||O|| = ||P_0 - X_0||$$

donde $X_0 = P_0 - O$ es un punto de la recta escogido de tal manera que O sea ortogonal a la recta.



Sea ahora $\tilde{N}$ *cualquier* vector (no nulo) ortogonal a N. Entonces, puesto que $O = P_0 - X_0 = P_0 - (P + t_0 N)$ para cierto escalar t_0, entonces

$$O \cdot \tilde{N} = (P_0 - (P + t_0 N)) \cdot \tilde{N} = (P_0 - P) \cdot \tilde{N}$$

y así, puesto que O y $\tilde{N}$ son paralelos, entonces

$$d = ||O|| = \frac{|(P_0 - P) \cdot \tilde{N}|}{||\tilde{N}||} \quad \textit{(fórmula de distancia)}$$

•) Por ejemplo, si la recta está en el plano y tiene la forma clásica $Ax+By+C = 0$, tome $\tilde{N} = (A, B)$, $P_0 = (x_0, y_0)$, $P = (x_1, y_1)$, y pruebe que

$$d = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}$$

•) Si la recta está en el espacio, y en su forma clásica es $\dfrac{x-p}{a} = \dfrac{y-q}{b} = \dfrac{z-r}{c}$, escoja $\tilde{N}$ ortogonal a (a, b, c), $P = (p, q, r)$, y $P_0 = (x_0, y_0, z_0)$ para obtener

$$d = \frac{|(x_0 - p, y_0 - q, z_0 - r) \cdot \tilde{N}|}{||\tilde{N}||}$$

¿Cuál puede ser $\tilde{N}$?

b) Tomemos ahora un plano en el espacio de la forma estándar $(X - P) \cdot N = 0$ y un punto P_0. Nuevamente, llegamos a que la distancia del punto P_0 a la recta estará dada por

$$d = ||O|| = ||P_0 - X_0||$$

donde $X_0 = P_0 - O$ es un punto del plano escogido de tal forma que O, pasando por X_0, sea ortogonal al plano (¿podría el lector hacer una figura que ilustre esto?). Ahora: puesto que $O = P_0 - X_0$ entonces

$$\begin{aligned} O \cdot N &= P_0 \cdot N - X_0 \cdot N = P_0 \cdot N - P \cdot N \\ &= (P_0 - P) \cdot N \end{aligned}$$

y así, regresaremos a una ecuación para la distancia muy similar a la que obtuvimos para las rectas en a):

$$d = ||O|| = \frac{|(P_0 - P) \cdot N|}{||N||} \quad \textit{(fórmula de distancia)}$$

•) Por ejemplo, si el plano tiene la forma clásica $Ax + By + Cz + D = 0$, tome $N = (A, B, C)$, $P_0 = (x_0, y_0, z_0)$, $P = (x_1, y_1, z_1)$, y pruebe que

$$d = \frac{|Ax_0 + By_0 + Cz_0 + D|}{\sqrt{A^2 + B^2 + C^2}}$$

c) Ilustre lo analizado en a) y b) mediante ejemplos concretos de rectas (en el plano y en el espacio), y planos en el espacio.

35) Basándose en la fórmula anterior de distancia entre un punto y un plano, encuentre la distancia al origen $(0,0,0)$ del plano $Ax+By+Cz+D=0$.

36) a) En una economía Walras-Cassel con tres mercancías y dos factores de producción, se tiene que

$$A = \begin{bmatrix} 1 & 2 & 5 \\ \frac{1}{4} & 3 & 1 \end{bmatrix} \quad , \quad r = \begin{bmatrix} 7 \\ 12 \end{bmatrix}$$

¿Cuál es el vector x^T de cantidades producidas?

b) Si $p = (0.2,\ 0.3,\ 0.5)^T$ es el vector de precios de las mercancías, ¿cuál será el vector de precios v de los factores?

c) Pruebe *la ley de Walras*: $px = vr$.

37) ¿Cuál cree usted que es la razón (económica) por la que el sistema Walras-Cassel

$$\begin{array}{lllll} x_1 + 2x_2 = 10 & ; & x_1 = \dfrac{10}{p_1} & ; & v_1 + 4v_2 = p_1 \\ 4x_1 + 5x_2 = 30 & ; & x_2 = \dfrac{1}{p_2} & ; & 2v_1 + 5v_2 = p_2 \end{array}$$

no tiene solución económicamente significativa?

*38) Encuentre (si existen) posibles recursos r_1, r_2 tales que el sistema

$$\begin{aligned} &x_1 + 2x_2 = r_1 \\ &4x_1 + 5x_2 = r_2 \\ &x_1 = \frac{8}{p_1}, \quad x_2 = \frac{10}{p_2} \\ &v_1 + 4v_2 = p_1 \\ &2v_1 + 5v_2 = p_2 \end{aligned}$$

esté en equilibrio. ¿Podrían no ser únicos estos coeficientes?

Lección 5

Bases y dimensión

Introducción

Aunque el uso de los vectores en el plano, sus operaciones de suma y producto escalar, y su relación con los números complejos (Volumen 0 (Fundamentos)), ya eran conocidos por los matemáticos en 1830, la creación en 1843 de un análogo espacial de los números complejos debida a William R. Hamilton [1805–1865] sorprendió a todos. La razón era que hasta ese momento *todos* los números que conocían los matemáticos poseían la propiedad conmutativa de la multiplicación (es decir, $ab = ba$) y Hamilton encontró que sus números, conocidos como *cuaterniones*, no satisfacían esta propiedad.

Recordemos que los números complejos son objetos de la forma $a + bi$, donde a y b son números reales y $i = \sqrt{-1}$. Los cuaterniones, a su vez, son números de la forma

$$a + bi + cj + dk$$

donde a, b, c, d son números reales; i, j y k poseen la misma propiedad de $\sqrt{-1}$, es decir,

$$i^2 = j^2 = k^2 = -1$$

Claramente, los cuaterniones se pueden sumar de manera natural. Sin embargo, para la multiplicación, Hamilton tuvo que especificar los productos de i con j, i con k y j con k. Los definió así:

$$jk = i, \quad kj = -i, \quad ki = j, \quad ik = -j, \quad ij = k, \quad ji = -k$$

(¿Reconoce el lector este tipo de multiplicación? Recuerde el producto cruz de vectores en $\mathbb{R}^3$). Con esto, los cuaterniones no satisfacen la propiedad conmutativa para el producto, y sin embargo comparten con los números reales muchas otras de sus propiedades. Y aunque estos números no resultaron ser de tanta utilidad como Hamilton esperaba, pudo, aun así, aplicarlos a un gran número de problemas físicos y geométricos. Por otra parte, no mucho después de la creación de los cuaterniones, el famoso matemático Arthur Cayley introdujo en 1858, de forma general, *las matrices*. Estas, ya sabemos, aunque están sujetas a las operaciones usuales del álgebra, tampoco satisfacen la propiedad conmutativa de la multiplicación. Incluso, el producto de dos matrices puede ser cero sin que ninguna de ellas lo sea.

Los cuaterniones y las matrices fueron los heraldos de una larga lista de nuevas estructuras algebraicas con más y más extrañas propiedades. Y aunque la creación de estos sistemas para propósitos particulares retaba las verdades establecidas de la aritmética ordinaria, no tuvieron un impacto directo sobre el análisis: los números reales y complejos se utilizaban con propósitos diferentes en donde su aplicabilidad era incuestionable. Aun así, la física señalaba que la aritmética ordinaria no podía describir muchas situaciones en las que realmente tenía una aplicación muy limitada. *Por primera vez, los matemáticos estaban llegando a la conclusión de que la verdad absoluta de las matemáticas que la Grecia clásica intentaba garantizar partiendo de "verdades evidentes por sí mismas" y de un proceso deductivo, simplemente no existía.*

Por todo esto, el desarrollo que distingue a las matemáticas del siglo XX es el surgimiento de una disciplina cuyo concepto central no es ni el número ni la cantidad sino la *estructura*. La emersión y difusión de estructuras en el siglo XX (tales como la de espacio vectorial que estudiaremos en esta lección) tiene dos causas fundamentales. Por una parte, era deseable interpretar el copioso material acumulado hasta el siglo XIX en distintas y, a menudo, aparentemente inconexas ramas de las matemáticas, puesto que esto posibilitaba la comprensión de muchos resultados de esos campos desde un punto vista unificado y promovía la creación de otros nuevos; por otro lado, el estudio de problemas matemáticos relacionados con la física (la mecánica cuántica, por ejemplo) se convirtió en un hecho crucial en el desarrollo posterior de diferentes estructuras matemáticas.

1. Definición de espacio vectorial

Ya habíamos mencionado que una característica esencial del espacio n-dimensional es la existencia de las operaciones de adición y de multiplicación por escalar, cuyas propiedades nos recuerdan las operaciones con números (asociatividad, conmutatividad, etc.). Sin embargo, no sólo los vectores tienen estas características. También el conjunto de las matrices poseen igual estructura a la de las cantidades vectoriales físicas; y otros ejemplos aparecerán en el desarrollo de la presente lección, mostrando que hechos esenciales de muchos conjuntos aparentemente disímiles se pueden resumir en unas cuantas propiedades que les determinan su *estructura básica*[1], y así se hace posible aplicar los métodos del álgebra lineal a un rango muy amplio de problemas en ciencia teórica.

Definición 1. (Espacio vectorial (Grassmann(1846), Peano (1888))) Consideremos un conjunto no vacío V con dos operaciones (suma y multiplicación por escalar) que satisfagan:

i) *Adición*: a cada par x, $y \in V$ se le asigna un único $x + y \in V$, llamado *suma*.

ii) *Multiplicación por escalar*: a cada $x \in V$ y a cada escalar (número real) $k \in \mathbb{R}$ se les asigna un único $kx \in V$ llamado *producto por escalar*.

Llamamos a V un *espacio vectorial* (y a sus elementos *vectores*) si las siguientes afirmaciones se cumplen para todo x, y, $z \in V$:

a) $x + y = y + x$ (ley conmutativa)

b) $(x + y) + z = x + (y + z)$ (ley asociativa)

c) Existe un vector $0 \in V$, que llamaremos el *vector cero*, tal que $x+0 = 0 + x = x$ para cualquier $x \in V$.

d) Para cada $x \in V$ existe un vector en V, que denotaremos por $-x$, tal que $x + (-x) = (-x) + x = 0$. Al vector $-x$ lo llamaremos *inverso aditivo de* x y escribiremos la suma del vector x y del inverso aditivo de y así:

$$(x + (-y)) = x - y$$

e) $l(kx) = (lk)x$ para todo par de escalares $l, k \in \mathbb{R}$.

f) $k(x + y) = kx + ky$ para cada escalar $k \in \mathbb{R}$.

g) $(k + l)x = kx + lx$ para cualquier par de escalares $k, l \in \mathbb{R}$.

h) $1x = x$. [2]

A partir de estos ocho axiomas podemos comenzar a deducir diversas propiedades de un espacio vectorial. Algunas de las más “obvias” son las siguientes:

[1] Es clara aquí la influencia de la estructura axiomática de los *Elementos* de Euclides (volumen 0 (Fundamentos)).

[2] Se ha visto necesario incluir este axioma aparentemente trivial pues no es posible deducirlo de los otros siete axiomas que definen a un espacio vectorial.

Teorema 1.
Sea V un espacio vectorial. Entonces

a) $k0 = 0$ *para todo* $k \in \mathbb{R}$.

b) $0x = 0$ *para todo* $x \in V$.

c) $(-1)x = -x$ *para todo* $x \in V$.

Demostración

a) Las condiciones c) y f) de la definición 1 implican que $k0 = k(0+0) = k0 + k0$. Sumando $-k0$ a ambos lados de esta igualdad y utilizando las condiciones b), c) y d) de la definición 1 se tiene que

$$\begin{aligned} k0 + (-k0) &= k0 + k0 + (-k0) \\ 0 &= k0 + (k0 + (-k0)) \\ 0 &= k0 + 0 \\ 0 &= k0 \end{aligned}$$

b) Ya que $0 + 0 = 0$, entonces la condición g) de la definición 1 implica que $0x = (0+0)x = 0x + 0x$. Sumando $-0x$ a ambos lados de esta igualdad y utilizando las condiciones b), c) y d) de la definición 1 se tiene que

$$\begin{aligned} 0x + (-0x) &= 0x + 0x + (-0x) \\ 0 &= 0x + [0x + (-0x)] \\ 0 &= 0x + 0 \\ 0 &= 0x \end{aligned}$$

c) Utilizando la parte b) de este teorema se tiene que

$$0 = 0x = (1 + (-1))x = 1x + (-1)x = x + (-1)x$$

Sumando $-x$ a ambos lados de esta igualdad y utilizando las condiciones a), c) y d) de la definición 1 se tiene que

$$\begin{aligned} -x &= 0 + (-x) = x + (-1)x + (-x) \\ &= x + (-x) + (-1)x = [x + (-x)] + (-1)x \\ &= 0 + (-1)x = (-1)x \qquad \blacksquare \end{aligned}$$

Y ahora presentamos numerosos ejemplos (de índole matemática diversa) que satisfacen, todos, los ocho axiomas que definen a un espacio vectorial.

Ejemplo 1. (Las matrices forman un espacio vectorial)
El conjunto de las matrices $m \times n$, $\mathfrak{M}_{m\times n}$, es un espacio vectorial, y el lector podría comprobarlo utilizando los resultados de la lección 2.

Ejemplo 2. (Los vectores forman un espacio vectorial)
El conjunto de vectores $\mathbb{R}^n$ es también un espacio vectorial, como el lector puede fácilmente comprobarlo utilizando los resultados de la lección 4.

Ejemplo 3. (Soluciones del sistema homogéneo como espacio vectorial)
Sea A una matriz $m \times n$. El *conjunto W de las soluciones* del sistema lineal homogéneo $AX = 0$ es un espacio vectorial.

Solución

Para comprobar que el conjunto-solución del sistema $AX = 0$, es un espacio vectorial en *sí mismo*, primero debemos reconocer que es un subconjunto de las matrices $n \times 1$ y que, por tanto, hereda ciertas propiedades de este conjunto. Revisemos entonces, propiedad tras propiedad, la definición 1 de espacio vectorial y comprobemos que nuestro conjunto las satisface:

a) En primer lugar, observemos que si $x, y \in W$, entonces $Ax = 0$ y $Ay = 0$; luego $A(x + y) = Ax + Ay = 0 + 0 = 0$; y así $x + y \in W$.

b) En segundo lugar, si $x \in W$ y $k \in \mathbb{R}$ es un escalar, entonces, como $Ax = 0$, se tendrá también que $A(kx) = kAx = k0 = 0$.

Observamos de la condición ii) arriba, que para $k = 0$ y $k = -1$ se tiene que si $x \in W$, entonces $0x = 0 \in W$ y $(-1)x = -x \in W$. Por consiguiente, W satisface las condiciones c) y d) de espacio vectorial; y las otras propiedades: a), b), e), f), g), y h) son heredadas del conjunto de las matrices $n \times 1$. Luego W satisface *todas* las propiedades de espacio vectorial. ▲

Ejemplo 4. (¿Cuáles hiperplanos en $\mathbb{R}^n$ son espacios vectoriales?)
El hiperplano en $\mathbb{R}^n$,

$$H = \{ (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \,/\, k_1x_1 + k_2x_2 + \cdots + k_nx_n = 0 \}$$

donde las k_i son constantes fijas, es un espacio vectorial. Así, las rectas y los planos que *pasan por el origen* (pues $0 = (0, \ldots, 0)$ pertenece al conjunto H) también son espacios vectoriales.

Solución

De manera similar a lo efectuado en el ejemplo 3, inmediatamente observamos que como éste también es un subconjunto de $\mathbb{R}^n$, sólo es necesario probar dos puntos:

a) Si $x,\, y \in H$, entonces $x + y \in H$.

b) Si $x \in H$ y $k \in \mathbb{R}$, $kx \in H$.

Y que esto último es así, se ve enseguida:

a) Si $x = (x_1, \ldots, x_n) \in H$ y $y = (y_1, \ldots, y_n) \in H$, entonces $k_1x_1 + k_2x_2 + \cdots + k_nx_n = 0$ y $k_1y_1 + k_2y_2 + \cdots + k_ny_n = 0$; y así, sumando ambas ecuaciones término a término, se obtiene $k_1(x_1 + y_1) + k_2(x_2 + y_2) + \cdots + k_n(x_n + y_n) = 0$; es decir, $(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \in H$ o $x + y \in H$.

b) Si $x = (x_1, \ldots, x_n) \in H$ y $k \in \mathbb{R}$, entonces, como $k_1x_1 + k_2x_2 + \cdots + k_nx_n = 0$, se tendrá que $kk_1x_1 + kk_2x_2 + \cdots + kk_nx_n = 0$ o $k_1(kx_1) + k_2(kx_2) + \cdots + k_n(kx_n) = 0$ o, lo que es equivalente,

$$(kx_1, kx_2, \ldots, kx_n) = k(x_1, \ldots, x_n) \in H$$

Nota 1.
El conjunto de soluciones del sistema lineal no-homogéneo $AX = b$, $b \neq 0$ *no es un espacio vectorial*. De forma similar, los hiperplanos $H = \{ (x_1, \ldots, x_n) \in \mathbb{R}^n \,/\, k_1x_1 + \cdots + k_nx_n = b \}$, $b \neq 0$, que *no pasan por el origen*, tampoco son espacios vectoriales. ¿Podría el lector indicar por qué?; es decir, ¿qué propiedades de un espacio vectorial, definitivamente, no pueden satisfacer?

Ejemplo 5. (Un hiperplano particular)
Consideremos el vector $(1,1) \in \mathbb{R}^2$ y el conjunto $\mathcal{L} = \{ k(1,1) \,|\, k \in \mathbb{R} \}$. Con lo realizado en los ejemplos anteriores ahora es fácil observar que este conjunto es un espacio vectorial que es subconjunto del espacio vectorial $\mathbb{R}^2$. Si dibujamos $\mathcal{L}$ sobre el plano $\mathbb{R}^2$, notamos que $\mathcal{L}$ forma una línea que pasa por el origen (cuando $k = 0$) y por el punto $(1,1)$ (cuando $k = 1$). En general, con todo vector $(x_0, y_0) \in \mathbb{R}^2$ se puede formar el espacio vectorial $\mathcal{L} = \{k(x_0, y_0) \,|\, k \in \mathbb{R}\}$; si $(x_0, y_0) \neq (0,0)$ el conjunto $\mathcal{L}$ es una recta que pasa por el origen y por (x_0, y_0), como en la figura 1; si $(x_0, y_0) = (0,0)$, entonces $\mathcal{L} = \{(0,0)\}$.

Notemos, además, que si $(x, y) \in \mathcal{L}$, entonces $x = kx_0$, $y = ky_0$ para cierto $k \in \mathbb{R}$. Si $x_0 \neq 0$, entonces $\dfrac{x}{x_0} = k$ y así $y = \dfrac{y_0}{x_0}x$ que es la forma cartesiana de la recta $\mathcal{L}$. También podríamos escribirla como $x_0y - y_0x = 0$ y así aparecer como un *hiperplano* de los definidos en el ejemplo 4.

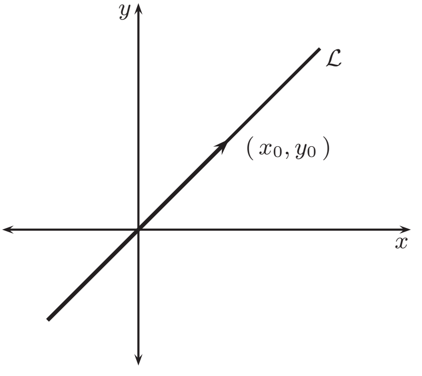


Figura 1

Ejemplo 6. (Las matrices simétricas como un espacio vectorial)
El conjunto de todas las matrices *simétricas* de orden $n \times n$ es un espacio vectorial subconjunto del espacio vectorial $\mathfrak{M}_{n\times n}$:

a) Observemos que si A y B son matrices simétricas, entonces la matriz $A + B$ también es simétrica:

$$(A + B)^T = A^T + B^T = A + B$$

b) Además, para todo $k \in \mathbb{R}$, kA también es una matriz simétrica:

$$(kA)^T = kA^T = kA$$

Con esto hemos probado que el conjunto de las matrices simétricas satisface las condiciones iniciales i) y ii) para la adición y la multiplicación por escalar de la definición 1. Las otras propiedades de la definición 1 de espacio vectorial deben cumplirse ya que corresponden a propiedades generales del espacio de matrices $\mathfrak{M}_{n\times n}$, del cual el conjunto de las matrices simétricas es un subconjunto.

Ejemplo 7. (Los números complejos forman un espacio vectorial)
El conjunto de los números complejos $\mathbb{C} = \{ a + ib \,/\, a,\ b \in \mathbb{R},\ i^2 = -1 \}$ es un espacio vectorial. ¿Puede el lector comprobarlo?

Ejemplo 8. (Las funciones $f : S(\subseteq \mathbb{R}) \longrightarrow \mathbb{R}$ forman un espacio vectorial)
Sea $\mathcal{F}_S$ el conjunto de las funciones de un subconjunto no vacío $S \subseteq \mathbb{R}$ en el conjunto de los números reales; es decir, $\mathcal{F}_S = \{f : S \longrightarrow \mathbb{R}/f(\cdot) \text{ es una función}\}$. Definiendo como 0 la función que asigna a cada $x \in S$ el número cero y $(-f)(x) = -f(x)$, se puede observar que $\mathcal{F}_S$ es un espacio vectorial. ¿Podría el lector mostrar esto?

Ejemplo 9. (Los polinomios forman un espacio vectorial)
Mostrar que el conjunto de todos los polinomios de la forma $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ es un espacio vectorial por sí mismo y que es subconjunto del espacio vectorial de todas las funciones sobre $\mathbb{R}$, $\mathcal{F}_{\mathbb{R}}$, es tarea sugerida para el lector.

Ejemplo 10. (Un caso particular)
P_n, el conjunto de todos los polinomios de grado *menor o igual* a n ($n \in \mathbb{N}$ fijo) es, a la vez, un espacio vectorial subconjunto del espacio vectorial de *todos* (sin importar el grado) los polinomios de la forma $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$.

a. Combinaciones lineales

Una operación mixta que es inmediatamente realizable en un espacio vectorial es la *combinación lineal* de vectores; de hecho, esta es la *operación esencial y central* del álgebra lineal.

Definición 2. (Combinación lineal)
Sea V un espacio vectorial y $x \in V$. Diremos que x es una *combinación lineal de los vectores* $x_1, x_2, \ldots, x_m \in V$ si existe una colección de escalares $k_1, k_2, \ldots, k_m \in \mathbb{R}$ tales que

$$x = k_1 x_1 + k_2 x_2 + \cdots + k_m x_m$$

Ejemplo 11. (Combinaciones lineales en $\mathbb{R}^2$)
Consideremos el vector $(2, -1)$. Veamos que este vector puede ser expresado como una combinación lineal de los vectores $(1, 2)$ y $(3, 3)$.

Solución

En efecto, pues basta encontrar $k_1, k_2 \in \mathbb{R}$ tales que

$$(2, -1) = k_1(1, 2) + k_2(3, 3)$$

es decir,

$$\begin{aligned} 2 &= k_1 + 3k_2 \\ -1 &= 2k_1 + 3k_2 \end{aligned}$$

y resolviendo este sistema de ecuaciones lineales tenemos que $k_1 = -3$ y $k_2 = \frac{5}{3}$ o, equivalentemente, $(2, -1) = -3(1, 2) + \frac{5}{3}(3, 3)$ (figura 2).

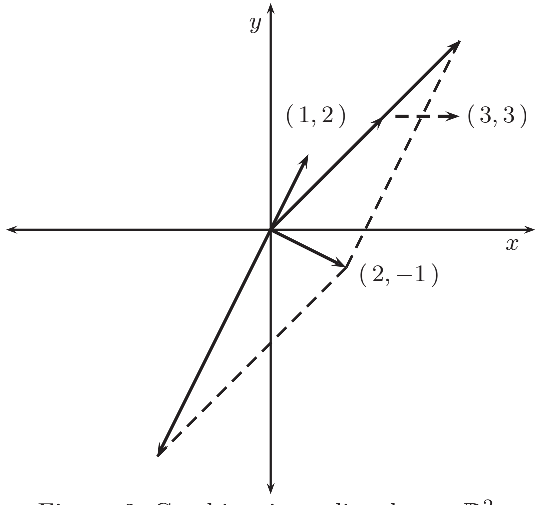


Figura 2. Combinaciones lineales en $\mathbb{R}^2$

Ejemplo 12. (Combinaciones lineales en $\mathbb{R}^3$)
Mostremos que el vector $x = (3, -2, 7)$ es una combinación lineal de los vectores $x_1 = (1, 4, 5)$, $x_2 = (2, 1, 3)$ y $x_3 = (2, -2, 1)$.

Solución
Debemos encontrar unos escalares k_1, k_2 y k_3 no todos nulos tales que $x = k_1x_1 + k_2x_2 + k_3x_3$. Es decir, encontrar k_1, k_2, k_3 tales que satisfagan el sistema lineal

$$\begin{aligned} 3 &= k_1 + 2k_2 + 2k_3 \\ -2 &= 4k_1 + k_2 - 2k_3 \\ 7 &= 5k_1 + 3k_2 + k_3 \end{aligned}$$

La matriz aumentada de este sistema de ecuaciones lineales es:

$$\left[\begin{array}{rrrcr} 1 & 2 & 2 & | & 3 \\ 4 & 1 & -2 & | & -2 \\ 5 & 3 & 1 & | & 7 \end{array}\right] \begin{array}{l} F_1 \\ F_2 \\ F_3 \end{array}$$

Aplicando el método de eliminación gaussiana se tiene que:

$$\left[\begin{array}{rrrcr} 1 & 2 & 2 & | & 3 \\ 0 & -7 & -10 & | & -14 \\ 0 & -7 & -9 & | & -8 \end{array}\right] \begin{array}{l} \\ F_2 \longleftrightarrow F_2 - 4F_1 \\ F_3 \longleftrightarrow F_3 - 5F_1 \end{array}$$

$$\left[\begin{array}{rrrcr} 1 & 2 & 2 & | & 3 \\ 0 & 1 & \frac{10}{7} & | & 2 \\ 0 & -7 & -9 & | & -8 \end{array}\right] \quad F_2 \longleftrightarrow -\tfrac{1}{7}F_2$$

$$\begin{bmatrix} 1 & 0 & -\frac{6}{7} & | & -1 \\ 0 & 1 & \frac{10}{7} & | & 2 \\ 0 & 0 & 1 & | & 6 \end{bmatrix} \quad \begin{array}{l} F_1 \longleftrightarrow F_1 - 2F_2 \\ \\ F_3 \longleftrightarrow F_3 + 7F_2 \end{array}$$

$$\begin{bmatrix} 1 & 0 & 0 & | & \frac{29}{7} \\ 0 & 1 & 0 & | & -\frac{46}{7} \\ 0 & 0 & 1 & | & 6 \end{bmatrix} \quad \begin{array}{l} F_1 \longleftrightarrow F_1 + \frac{6}{7}F_3 \\ F_2 \longleftrightarrow F_2 - \frac{10}{7}F_3 \end{array}$$

Así, $k_1 = \frac{29}{7}$, $k_2 = -\frac{46}{7}$ y $k_3 = 6$. Por tanto, $(3, -2, 7) = \frac{29}{7}(1, 4, 5) - \frac{46}{7}(2, 1, 3) + 6(2, -2, 1)$.

Ejemplo 13. (Combinaciones lineales en $\mathbb{R}^4$)
Consideremos el vector $x = (3, 2, 2, 0)$. Expresemos este vector como una combinación lineal de los vectores $x_1 = (1, 1, 1, 0)$, $x_2 = (-1, 0, 1, 1)$ y $x_3 = (-2, 1, 0, -1)$.

Solución
Necesitamos encontrar tres escalares k_1, k_2 y k_3 no todos nulos tales que

$$\begin{bmatrix} 3 \\ 2 \\ 2 \\ 0 \end{bmatrix} = k_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + k_2 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix} + k_3 \begin{bmatrix} -2 \\ 1 \\ 0 \\ -1 \end{bmatrix}$$

y esto es equivalente al sistema de ecuaciones

$$\begin{aligned} 3 &= k_1 - k_2 - 2k_3 \\ 2 &= k_1 + k_3 \\ 2 &= k_1 + k_2 \\ 0 &= k_2 - k_3 \end{aligned}$$

Resolviendo este sistema de ecuaciones lineales se tiene que $k_1 = \frac{9}{4}$, $k_2 = k_3 = -\frac{1}{4}$.

Ejemplo 14. (Combinaciones lineales en el espacio $\mathfrak{M}_{3\times 4}$)
La matriz

$$A = \begin{bmatrix} 3 & -1 & -17 & -9 \\ -5 & 3 & -18 & -27 \\ -16 & 8 & -46 & -30 \end{bmatrix}$$

es una combinación lineal de las matrices

$$B = \begin{bmatrix} 1 & 3 & -4 & 2 \\ 5 & 6 & -1 & 1 \\ 3 & 1 & -2 & 5 \end{bmatrix} \quad y \quad C = \begin{bmatrix} 0 & 2 & 1 & 3 \\ 4 & 3 & 3 & 6 \\ 5 & -1 & 8 & 9 \end{bmatrix}$$

ya que $A = 3B + (-5)C$.

Ejemplo 15. (Combinaciones lineales en el espacio de polinomios P_n)
Sabemos que un polinomio con coeficientes reales de grado menor o igual que n es una función de la forma

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

donde $a_i \in \mathbb{R}$, para $i = 0, 1, \ldots, n$. En este espacio vectorial, el polinomio $p(x) = 3x^4 + 2x^3 + 8x^2 + 9x + 1$ puede expresarse como una combinación lineal de los polinomios $p_1(x) = x^3 + x^2 + 2$ y $p_2(x) = x^4 + 2x^2 + 3x - 1$ ya que $p(x) = 2p_1(x) + 3p_2(x)$.

Ejemplo 16. (Combinaciones lineales en el espacio de funciones $\mathcal{F}_{\mathbb{R}}$)
La función

$$f(x) = 2e^x + x + 3x^2 + 5$$

es una combinación lineal de $f_1(x) = e^x$, $f_2(x) = 1$, $f_3(x) = x$ y $f_4(x) = x^2$, pues

$$f(x) = 2f_1(x) + f_3(x) + 3f_4(x) + 5f_2(x)$$

b. Subespacios vectoriales

Informalmente, un *subespacio vectorial es un espacio vectorial contenido dentro de otro espacio vectorial.* Pero lo importante aquí, como veremos, es que la relación entre un espacio vectorial y sus subespacios nos puede indicar información lineal adicional.

Definición 3. (Subespacio vectorial)
Sean V un espacio vectorial y W un subconjunto no vacío de V. Decimos que W es un *subespacio vectorial de* V si W es también un espacio vectorial con las mismas operaciones de suma y multiplicación por escalar de V.

Puesto que $W \subseteq V$, podría ser claro que para verificar si W es un subespacio de V, no es necesario comprobar que todas las ocho propiedades de la definición 1 se satisfacen. Basta determinar si la suma y el producto por escalar son *operaciones cerradas* en W; es decir, basta comprobar que para todo par de elementos $x, y \in W$ y todo escalar $k \in \mathbb{R}$ se cumpla que:

a) $x + y \in W$ b) $kx \in W$

Las otras propiedades se tendrán inmediatamente por el hecho de que W "hereda" éstas del espacio vectorial V.

Nota 2.
Podemos, además, observar que:

a) Todo subespacio W de un espacio vectorial V contiene al vector 0 (cero); en efecto, como W es no vacío, sea $x \in W$ y $k = 0$, entonces $kx = 0x = 0 \in W$ de acuerdo con la condición b) de arriba.

b) Todo espacio vectorial V tiene por lo menos dos subespacios (que llamaremos *subespacios triviales*): el conjunto cuyo único elemento es el vector 0 (cero) del espacio, y el conjunto formado por todo el espacio V; es decir, $W = \{0\}$ y $W = V$ son los subespacios triviales de V.

Ejemplo 17. (¿Cuáles son los subespacios del espacio vectorial $\mathbb{R}$?)
Mostremos que los únicos subespacios del espacio vectorial $\mathbb{R}$ son los triviales: $\{0\}$ y $\mathbb{R}$.

Solución

Sea $W \neq \{0\}$ un subespacio de $\mathbb{R}$. Entonces existe un número l diferente de cero en W. Esto implica que $1 = (\frac{1}{l})l \in W$ y así $k = k1 \in W$ para todo $k \in \mathbb{R}$. Por tanto, $W = \mathbb{R}$.

Ejemplo 18. (¿Cuáles son los subespacios del espacio vectorial $\mathbb{R}^2$?)
En $\mathbb{R}^2$ los únicos subespacios son $\{(0,0)\}$, $\mathbb{R}^2$, y todos los múltiplos por escalar de algún vector no-nulo en $\mathbb{R}^2$ (recordemos que, geométricamente, estos últimos son las rectas a través del origen $(0,0)$). Para comprobarlo, supongamos que W es un subespacio no-trivial de $\mathbb{R}^2$ (es decir, $W \neq \{(0,0)\}$ y $W \neq \mathbb{R}^2$). Sea $(x_0, y_0) \in W$, $(x_0, y_0) \neq (0,0)$. Entonces, para todo $k \in \mathbb{R}$, $k(x_0, y_0) \in W$ y, por tanto, el subespacio $\{k(x_0, y_0) \,/\, k \in \mathbb{R}\}$ de $\mathbb{R}^2$ está incluido en W. Para ver que, de hecho, *coincide* con W, sea $(x_1, y_1) \in W$ pero $(x_1, y_1) \neq l(x_0, y_0)$ para todo l; entonces, si $(x, y) \in \mathbb{R}^2$ es cualquiera y podemos encontrar escalares k_1, k_2 tales que

$$(x, y) = k_1(x_1, y_1) + k_2(x_0, y_0) \tag{1}$$

habremos demostrado que $W = \mathbb{R}^2$ lo que, por hipótesis, es una contradicción. Pero la ecuación (1) es equivalente a

$$\begin{aligned} x &= k_1x_1 + k_2x_0 \\ y &= k_1y_1 + k_2y_0 \end{aligned} \tag{2}$$

y el determinante $x_1y_0 - x_0y_1 \neq 0$, pues $x_1y_0 - x_0y_1 = 0$ implicaría $x_1 = lx_0$ y $y_1 = ly_0$ para algún $l \in \mathbb{R}$ y tendríamos que $(x_1, y_1) = l(x_0, y_0)$ y esto contradice nuestra hipótesis de arriba. Luego así se garantiza la existencia de k_1 y k_2 en (2) y, por ende, en (1).

Ejemplo 19. (¿Cuáles son los subespacios de $\mathbb{R}^3$?)
Se puede mostrar, de manera similar a lo realizado en el ejemplo anterior (aunque con un poco más de trabajo), que en $\mathbb{R}^3$ los subespacios son de tres tipos:

a) Los subespacios triviales: $\{(0,0,0)\}$ y $\mathbb{R}^3$.

b) Todas las rectas que pasan por el origen.

c) Todos los planos que pasan por el origen.

Ejemplo 20.
El conjunto de soluciones $x \in \mathbb{R}^n$ del sistema lineal homogéneo $Ax = 0$, donde $A \in \mathfrak{M}_{m\times n}$, es un subespacio vectorial de $\mathbb{R}^n$ (ejemplo 3).

Ejemplo 21.
El conjunto de polinomios de grado menor o igual m, P_m, es un subespacio vectorial del espacio P_n de polinomios de grado menor o igual que n para $0 < m \leq n$ (ejemplo 10).

Teorema 2. ***(Dos tipos generales de subespacios)***
Sea V un espacio vectorial y $W_1,\ W_2$ subespacios de V. Entonces

a) $W_1 + W_2 = \{x + y \,/\, x \in W_1,\, y \in W_2\}$ *es un subespacio de V.*

b) $W_1 \cap W_2$ *es un subespacio de V.*

Demostración

a) Probemos que W_1+W_2 es un subespacio: sean $x = w_1+v_1$ y $y = w_2+v_2$ dos elementos de $W_1 + W_2$, donde $w_1,\, w_2 \in W_1,\ v_1,\, v_2 \in W_2$. Entonces

   i) $x+y = (w_1+v_1)+(w_2+v_2) = (w_1+w_2)+(v_1+v_2) \in W_1+W_2$, pues $w_1 + w_2 \in W_1$ y $v_1 + v_2 \in W_2$.

   ii) Si $l \in \mathbb{R},\ lx = l(w_1 + v_1) = lw_1 + lv_1 \in W_1 + W_2,$ pues $lw_1 \in W_1$ y $lv_1 \in W_2$.

b) Probemos que $W_1 \cap W_2$ también es un subespacio:

   i) Si $x, y \in W_1 \cap W_2$, entonces $x, y \in W_1$ y $x, y \in W_2$; luego, $x+y \in W_1$ y $x + y \in W_2$; así, $x + y \in W_1 \cap W_2$.

   ii) Si $x \in W_1 \cap W_2$ y $l \in \mathbb{R}$, entonces $lx \in W_1$ y $lx \in W_2$ y, por tanto, $lx \in W_1 \cap W_2$. ■

Ejemplo 22.
Si consideramos los subespacios de $\mathbb{R}^2$:

$$\begin{aligned} W_1 &= \{\, m(1,0)\,/\, m \in \mathbb{R}\,\}\,(\equiv \text{ eje X}) \\ W_2 &= \{\, l(0,1)\,/\, l \in \mathbb{R}\,\}\,(\equiv \text{ eje Y}), \end{aligned}$$

entonces el lector debería convencerse, analítica y geométricamente, de que

a) $W_1 + W_2 = \mathbb{R}^2$; b) $W_1 \cap W_2 = \{\,(0,0)\,\}$

Definición 4. (Subespacio generado por vectores)
Sea V un espacio vectorial y $G \subseteq V$ un conjunto no vacío. Al conjunto de *todas* las posibles combinaciones lineales de elementos de G lo llamaremos el *subespacio generado por* G, y lo notaremos $\langle\, G\,\rangle$. Si $G = \{\,\beta_1, \beta_2, \ldots, \beta_n\,\}$, entonces diremos que $\langle\, G\,\rangle$ es el *subespacio generado por los vectores* β_1, β_2, $\ldots$, β_n. Obviamente, $\langle\, G\,\rangle$ es un subespacio de V como el lector fácilmente puede comprobar.

Ejemplo 23.

a) En $\mathbb{R}^2$, si $(x_0, y_0) \neq (0,0)$, entonces $\langle\,(x_0,y_0)\,\rangle = \{\, k(x_0,y_0)\,/\, k \in \mathbb{R}\,\} = \mathcal{L}$ es la línea recta dirigida por el vector (x_0, y_0) que pasa a través del origen.

b) En $\mathbb{R}^2$,

$$\begin{aligned} (1,0),(0,1) &= \{\, k(1,0) + l(0,1)\,/\, k,\; l \in \mathbb{R}\,\} \\ &= \{\,(k,l)\,/\, k, l \in \mathbb{R}\,\} = \mathbb{R}^2 \end{aligned}$$

c) En $\mathbb{R}^2$,

$$\begin{aligned} (1,2),(3,4) &= \{\, k(1,2) + l(3,4)\,/\, k, l \in \mathbb{R}\,\} \\ &= \{\,(k + 3l, 2k + 4l)\,/\, k, l \in \mathbb{R}\,\} = \mathbb{R}^2 \end{aligned}$$

pues si $(x,y) \in \mathbb{R}^2$ es dado, entonces

$$\begin{aligned} x &= k + 3l \\ y &= 2k + 4l \end{aligned}$$

donde, por la regla de Cramer,

$$k = \frac{\begin{vmatrix} x & 3 \\ y & 4 \end{vmatrix}}{\begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix}} = \frac{4x - 3y}{4 - 6} = \frac{3}{2}y - 2x \qquad l = \frac{\begin{vmatrix} 1 & x \\ 2 & y \end{vmatrix}}{\begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix}} = \frac{y - 2x}{4 - 6} = x - \frac{y}{2}$$

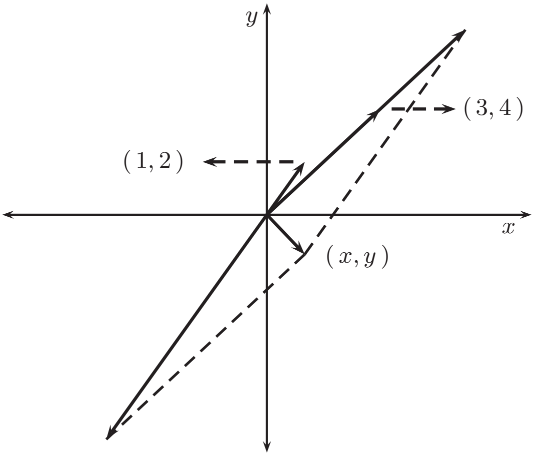


Figura 3

Observemos cómo *cualquier vector* (x, y) de $\mathbb{R}^2$ es una combinación lineal de *sólo dos vectores*: $(1, 2)$ y $(3, 4)$ (figura 3).

Ejemplo 24.
Si en $\mathbb{R}^3$ consideramos los planos

$$W_1 = \{ (x, y, z) \in \mathbb{R}^3 \,/\, 3x + 2y + 5z = 0 \}$$
$$W_2 = \{ (x, y, z) \in \mathbb{R}^3 \,/\, 2x + y - z = 0 \}$$

entonces sería conveniente en este punto que el lector (despejando x y colocándola en términos de y y z) se convenciera por sí mismo de que

$$W_1 = \langle (-\frac{2}{3}, 1, 0), (-\frac{5}{3}, 0, 1) \rangle$$
$$W_2 = \langle (-\frac{1}{2}, 1, 0), (\frac{1}{2}, 0, 1) \rangle$$

y también de que

$$W_1 + W_2 = \mathbb{R}^3$$

probando que al tomar cualquier vector $(a, b, c) \in \mathbb{R}^3$ éste se puede escribir como combinación lineal de los cuatro vectores $(-\frac{2}{3}, 1, 0), (-\frac{5}{3}, 0, 1)$, $(-\frac{1}{2}, 1, 0), (\frac{1}{2}, 0, 1)$, y resolviendo, por ejemplo, por el método gaussiano.

Ejemplo 25. (Espacio de soluciones de un sistema homogéneo)
Consideremos el siguiente sistema de ecuaciones lineales:

$$x_1 + x_2 - x_3 + 2x_4 = 0$$
$$3x_1 + 2x_2 + x_3 - x_4 = 0$$

La matriz aumentada para este sistema es

$$\left[\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 0 \\ 3 & 2 & 1 & -1 & 0 \end{array}\right] \quad \begin{array}{c} F_1 \\ F_2 \end{array}$$

$$\left[\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 0 \\ 0 & -1 & 4 & -7 & 0 \end{array}\right] \quad F_2 \longleftrightarrow F_2 - 3F_1$$

$$\left[\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 0 \\ 0 & 1 & -4 & 7 & 0 \end{array}\right] \quad F_2 \longleftrightarrow -F_2$$

$$\left[\begin{array}{cccc|c} 1 & 0 & 3 & -5 & 0 \\ 0 & 1 & -4 & 7 & 0 \end{array}\right] \quad F_1 \longleftrightarrow F_1 - F_2$$

Así, un sistema de ecuaciones equivalente al original es

$$x_1 + 3x_3 - 5x_4 = 0$$

$$x_2 - 4x_3 + 7x_4 = 0$$

Haciendo $x_3 = s$ y $x_4 = t$, para s, $t \in \mathbb{R}$, se tiene que

$$\begin{aligned} x_1 &= 5t - 3s \\ x_2 &= 4s - 7t; \end{aligned}$$

es decir,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = s \begin{bmatrix} -3 \\ 4 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 5 \\ -7 \\ 0 \\ 1 \end{bmatrix}$$

y así, el espacio-solución del anterior sistema lineal homogéneo es el subespacio de $\mathbb{R}^4$ generado $\langle(-3, 4, 1, 0), (5, -7, 0, 1)\rangle$.

Ejemplo 26. (Más sobre los sistemas lineales no-homogéneos)
Consideremos el sistema lineal $Ax = b$, donde $A = [a_{ij}]$ es una matriz $m \times n$, $x = (x_1, \ldots, x_n)^T$ y $b = (b_1, \ldots, b_m)^T$. Este sistema lineal puede escribirse como

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Por tanto, el vector b es una combinación lineal de las columnas de A. Así, *encontrar una solución del sistema $Ax = b$ es encontrar unos escalares $x_1, \ldots, x_n$ que permitan expresar a b como una combinación lineal de las columnas de A.*

Ejercicios 1

1) ¿Cuáles de los siguientes conjuntos son subespacios de $\mathbb{R}^n$?:

   a) $\{ (a_1, \ldots, a_n) \in \mathbb{R}^n \,/\, a_1 \geq 0 \}$
   b) $\{ (a_1, \ldots, a_n) \in \mathbb{R}^n \,/\, a_1 + 2a_2 = 5a_3 \}$
   c) $\{ (a_1, \ldots, a_n) \in \mathbb{R}^n \,/\, (a_1)^2 + (a_2)^2 + \cdots + (a_n)^2 = 1 \}$
   d) $\{ (a_1, \ldots, a_n) \in \mathbb{R}^n \,/\, a_1 + a_2 + \cdots + a_n = 1 \}$

2) En el espacio vectorial $\mathcal{F}_{\mathbb{R}}$ de todas las funciones $f : \mathbb{R} \longrightarrow \mathbb{R}$, ¿cuáles de los siguientes conjuntos son subespacios?:

   a) $\{ f : \mathbb{R} \longrightarrow \mathbb{R} \,/\, f(x) = f(-x) \text{ para todo } x \}$ (funciones pares)
   b) $\{ f : \mathbb{R} \longrightarrow \mathbb{R} \,/\, f(x) = -f(-x) \text{ para todo } x \}$ (funciones impares)
   c) $\{ f : \mathbb{R} \longrightarrow \mathbb{R} \,/\, f(1) = 0 \}$ (funciones que se anulan en $x = 1$)
   d) $\{ f : \mathbb{R} \longrightarrow \mathbb{R} \,/\, f(x) = ax + b \text{ para } a, b \in \mathbb{R} \text{ fijo} \}$ (funciones lineales)

3) En el espacio de las matrices $n \times n$, $\mathfrak{M}_{n\times n}$, ¿cuáles de los siguientes conjuntos son subespacios?:

   a) $\{ [a_{ij}] \in \mathfrak{M}_{n\times n} \,/\, a_{ij} = 0 \quad \text{si } i \neq j \}$ (matrices diagonales)
   b) $\{ [a_{ij}] \in \mathfrak{M}_{n\times n} \,/\, a_{ij} = 0 \quad \text{si } i \geq j \}$ (matrices triangulares superiores)
   c) $\{ [a_{ij}] \in \mathfrak{M}_{n\times n} \,/\, a_{ij} = 0 \quad \text{si } i \leq j \}$ (matrices triangulares inferiores)

4) Calcule explícitamente los siguientes subespacios:

   a) $\langle (1, 2, 3) \rangle$ en $\mathbb{R}^3$
   b) $\langle p_1(x), p_2(x) \rangle$ en P_3, donde $p_1(x) = 3 + x^2 + x^3$, $p_2(x) = x + 1$
   c) $\langle (1, 1, 1), (1, 2, 3), (0, 1, 7) \rangle$ en $\mathbb{R}^3$
   d) $\langle (0, 0) \rangle$ en $\mathbb{R}^2$
   e) $\langle (1, 1), (2, 2) \rangle$ en $\mathbb{R}^2$
   f) $\langle (1, 1), (3, 2) \rangle$ en $\mathbb{R}^2$
   g) $\langle (1, 2, 3) \rangle + \langle (1, 1, 1), (1, 0, 0) \rangle$ en $\mathbb{R}^3$
   h) $\langle 2 \rangle$ en $\mathbb{R}$

5) a) ¿Por qué será que, en general, la unión de dos espacios vectoriales no es, necesariamente, otro espacio vectorial? Dé ejemplos concretos de esto.

b) ¿Qué podemos decir del producto cartesiano de dos espacios vectoriales?

6) ¿Será que en $\mathbb{R}^3$ el vector $(1,-1,2)$ es combinación lineal de los vectores $(1,1,0)$, $(2,3,-1)$ y $(5,1,1)$?

7) ¿Será que en $\mathbb{R}^3$ el vector $(1,2,4)$ es combinación lineal de los vectores $(2,0,1)$ y $(-3,2,2)$?

8) ¿Será que en el espacio vectorial P_3 de todos los polinomios de grado ≤ 3, el polinomio

$$p(x) = x^3 + 2x + 1$$

es combinación lineal de los polinomios

$$p_1(x) = x, \quad p_2(x) = 3x + x^2, \quad p_3(x) = x^3 + x + 2 \ \ ?$$

9) En el espacio de las matrices $\mathfrak{M}_{3\times 3}$, ¿será que la matriz

$$A = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{4} \\ 1 & \frac{1}{2} & \frac{1}{2} \\ 1 & \frac{1}{3} & 2 \end{bmatrix}$$

es combinación lineal de

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_2 = A^{-1}, \ \ y \quad A_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}?$$

2. Las nociones de base y dimensión

Ya en varios ejemplos hemos apreciado que *cualquier* vector de cierto espacio puede expresarse como combinación lineal de unos *pocos vectores*. Reducir, si es posible, toda la *información lineal* de un espacio vectorial a unos *pocos vectores* es un paso esencial en el entendimiento de su estructura lineal interna. Para ello desarrollaremos los conceptos que aparecen a continuación.

a. Dependencia e independencia lineal

La noción geométrica sobre las condiciones bajo las cuales dos vectores son *colineales* (es decir, que están sobre la misma recta) o no, y las consecuencias que de esto se desprenden, es lo que nos conduce a la fundamental noción de *dependencia e independencia lineal.*

Definición 5. (Dependencia e independencia lineal)

a) Sea V un espacio vectorial y $\beta_1, \beta_2, \ldots, \beta_n \in V$. Diremos que $\beta = \{\beta_1, \beta_2, \ldots, \beta_n\}$ es un conjunto de vectores *linealmente dependientes* si existe un vector $\beta_i \in \beta$ que puede escribirse como combinación lineal de los restantes vectores en β; es decir, para ciertos escalares $k_1, k_2, \ldots, k_{i-1}$, $k_{i+1}, \cdots, k_n$, se tiene que

$$\beta_i = k_1\beta_1 + k_2\beta_2 + \cdots + k_{i-1}\beta_{i-1} + k_{i+1}\beta_{i+1} + \cdots + k_n\beta_n$$

o que

$$k_1\beta_1 + k_2\beta_2 + \cdots + k_{i-1}\beta_{i-1} - \beta_i + k_{i+1}\beta_{i+1} + \cdots + k_n\beta_n = 0$$

Esto es equivalente a decir que existe una colección de escalares k_1, $k_2, \ldots, k_n \in \mathbb{R}$ *no todos nulos* tales que

$$k_1\beta_1 + k_2\beta_2 + \cdots + k_n\beta_n = 0 \tag{1}$$

Veamos esto. Ya que no todos los escalares $k_1, k_2, \ldots, k_n$ son nulos, existe por lo menos uno distinto de cero. A este escalar lo podemos llamar k_i. Despejando de (1) se tiene que

$$\beta_i = -\frac{k_1}{k_i}\beta_1 - \cdots - \frac{k_{i-1}}{k_i}\beta_{i-1} - \frac{k_{i+1}}{k_i}\beta_{i+1} - \cdots - \frac{k_n}{k_i}\beta_n \tag{2}$$

Así, β_i es una combinación lineal de los otros elementos en β.

b) Un conjunto de vectores $\beta = \{\beta_1, \beta_2, \ldots, \beta_n\} \subseteq V$ es un conjunto de vectores *linealmente independientes* si no son linealmente dependientes; es decir, si *ninguno* de ellos puede expresarse como combinación lineal de los otros vectores. Esto significa que si existe una colección de escalares k_1, k_2, $\ldots$, $k_n \in \mathbb{R}$ tales que $k_1\beta_1 + k_2\beta_2 + \ldots + k_n\beta_n = 0$, entonces, necesariamente, $k_1 = k_2 = \ldots = k_n = 0$.

Nota 3.
Si $\beta = \{\beta_1, \beta_2, ..., \beta_n\}$ es linealmente dependiente también diremos que sus elementos $\beta_1, \beta_2, \ldots, \beta_n$ son linealmente dependientes. De similar manera para independencia lineal.

Ejemplo 27.
Sean $\beta_1 = (1, 2)$, $\beta_2 = (2, 0)$ y $\beta_3 = (4, 3)$ vectores de $\mathbb{R}^2$. Puesto que $\frac{3}{2}\beta_1 + \frac{5}{4}\beta_2 + (-1)\beta_3 = 0$, entonces β_1, β_2 y β_3 son linealmente dependientes en $\mathbb{R}^2$. En la figura 4 podemos observar que β_3 puede describirse como una combinación lineal de β_1 y β_2; de hecho, $\beta_3 = \frac{3}{2}\beta_1 + \frac{5}{4}\beta_2$.

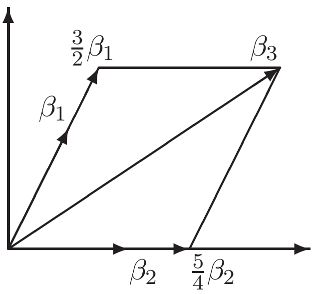


Figura 4

Ejemplo 28.
Sean $\beta_1 = (1,1,0)$, $\beta_2 = (2,3,4)$ y $\beta_3 = (4,5,4)$ tres vectores de $\mathbb{R}^3$. Puesto que

$$4\beta_1 + 2\beta_2 - 2\beta_3 = 4(1,1,0) + 2(2,3,4) - 2(4,5,4) = (0,0,0)$$

entonces β_1, β_2 y β_3 son linealmente dependientes.

Nota 4.
Cualquier conjunto de vectores de V que contenga al 0 es linealmente dependiente. ¿Podría el lector decir por qué?

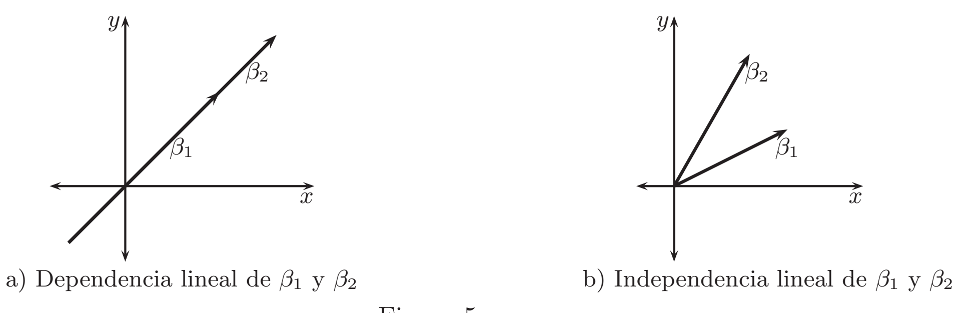


Figura 5

Ejemplo 29.
Decidamos si los cuatro vectores $(1,1,0,2)$, $(3,1,-1,4)$, $(5,0,-2,1)$, $(-1,-1,-1,-1)$ son o no linealmente independientes en $\mathbb{R}^4$.

Solución

Los vectores $(1,1,0,2)$, $(3,1,-1,4)$, $(5,0,-2,1)$, $(-1,-1,-1,-1)$ son linealmente independientes en $\mathbb{R}^4$ si, en el caso de que existan escalares k_1, k_2, k_3, k_4, tales que

$$k_1(1,1,0,2) + k_2(3,1,-1,4) + k_3(5,0,-2,1) + k_4(-1,-1,-1,-1) = 0,$$

se tendrá que $k_1 = k_2 = k_3 = k_4 = 0$. Ahora: esta última igualdad vectorial es equivalente a resolver el sistema homogéneo

$$\begin{aligned} k_1 + 3k_2 + 5k_3 - k_4 &= 0 \\ k_1 + k_2 - k_4 &= 0 \\ -k_2 - 2k_3 - k_4 &= 0 \\ 2k_1 + 4k_2 + k_3 - k_4 &= 0 \end{aligned}$$

La matriz aumentada de este sistema es

$$\begin{array}{cc} \left[\begin{array}{rrrr|r} 1 & 3 & 5 & -1 & 0 \\ 1 & 1 & 0 & -1 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ 2 & 4 & 1 & -1 & 0 \end{array}\right] & \begin{array}{l} F_1 \\ F_2 \\ F_3 \\ F_4 \end{array} \end{array}$$

Utilizando el método de eliminación gaussiana tenemos que

$$\left[\begin{array}{rrrr|r} 1 & 3 & 5 & -1 & 0 \\ 0 & 2 & 5 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & -2 & -9 & 1 & 0 \end{array}\right] \quad \begin{array}{l} \\ F_2 \longleftrightarrow F_1 - F_2 \\ \\ F_4 \longleftrightarrow F_4 - 2F_1 \end{array}$$

$$\left[\begin{array}{rrrr|r} 1 & 3 & 5 & -1 & 0 \\ 0 & 1 & \frac{5}{2} & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & -2 & -9 & 1 & 0 \end{array}\right] \quad \begin{array}{l} \\ F_2 \longleftrightarrow \frac{1}{2}F_2 \\ \\ \\ \end{array}$$

$$\left[\begin{array}{rrrr|r} 1 & 0 & -\frac{5}{2} & -1 & 0 \\ 0 & 1 & \frac{5}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & -1 & 0 \\ 0 & 0 & -4 & 1 & 0 \end{array}\right] \quad \begin{array}{l} F_1 \longleftrightarrow F_1 - 3F_2 \\ \\ F_3 \longleftrightarrow F_3 + F_2 \\ F_4 \longleftrightarrow F_4 + 2F_2 \end{array}$$

$$\left[\begin{array}{rrrr|r} 1 & 0 & -\frac{5}{2} & -1 & 0 \\ 0 & 1 & \frac{5}{2} & 0 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & -4 & 1 & 0 \end{array}\right] \quad \begin{array}{l} \\ \\ F_3 \longleftrightarrow 2F_3 \\ \\ \end{array}$$

$$\left[\begin{array}{rrrr|r} 1 & 0 & 0 & -6 & 0 \\ 0 & 1 & 0 & 5 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & -7 & 0 \end{array}\right] \quad \begin{array}{l} F_1 \longleftrightarrow F_1 + \frac{5}{2}F_3 \\ F_2 \longleftrightarrow F_2 - \frac{5}{2}F_3 \\ \\ F_4 \longleftrightarrow F_4 + 4F_3 \end{array}$$

$$\begin{bmatrix} 1 & 0 & 0 & -6 & | & 0 \\ 0 & 1 & 0 & 5 & | & 0 \\ 0 & 0 & 1 & -2 & | & 0 \\ 0 & 0 & 0 & 1 & | & 0 \end{bmatrix} \quad F_4 \longleftrightarrow -\tfrac{1}{7}F_4$$

Por tanto, $k_1 = k_2 = k_3 = k_4 = 0$ y así, los vectores $(1,1,0,2)$, $(3,1,-1,4)$, $(5,0,-2,1)$, $(-1,-1,-1,-1)$ sí son, efectivamente, linealmente independientes en $\mathbb{R}^4$.

Ejemplo 30.
Los vectores $\beta_1 = (1,1,0,1)$, $\beta_2 = (1,0,0,1)$, $\beta_3 = (1,-1,0,1)$ son linealmente dependientes en $\mathbb{R}^4$ ya que $\beta_1 - 2\beta_2 + \beta_3 = 0$.

Ejemplo 31. (Independencia lineal en las soluciones del sistema homogéneo)
Consideremos, de nuevo, el sistema lineal homogéneo $Ax = 0$, donde $A = [a_{ij}]$ es una matriz $m \times n$ y $x = (x_1, \ldots, x_n)^T$. Este sistema lineal puede escribirse también como

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

De acuerdo con la definición 5, este sistema tiene una solución $(x_1, \ldots, x_n)$ *no-nula* si, y sólo si, las columnas de la matriz A son *linealmente dependientes*. Si las columnas de A son linealmente independientes, la única solución es $x = (x_1, \ldots, x_n) = (0, \ldots, 0)$. ▲

Basados en lo anterior, podemos ahora dar la definición de uno de los conceptos centrales alrededor del cual gira todo el cuerpo del álgebra lineal:

Definición 6. (Base para un espacio vectorial (Peano (1888)))
Sea V un espacio vectorial cualquiera. Un conjunto $\beta \subseteq V$ es una *base* para V si satisface que:

a) β es un conjunto de vectores linealmente independiente en V; y

b) β genera V; es decir, $\langle \beta \rangle = V$.

Ejemplo 32. (Base canónica para $\mathbb{R}^n$)
En el espacio vectorial $\mathbb{R}^n$, el conjunto de vectores

$$\begin{aligned} e_1 &= (1, 0, 0, \ldots, 0) \\ e_2 &= (0, 1, 0, \ldots, 0) \\ &\vdots \\ e_n &= (0, 0, \ldots, 0, 1) \end{aligned}$$

forma una base para $\mathbb{R}^n$ (*base canónica*).

Solución
Los vectores $e_1, e_2, \cdots, e_n$ son linealmente independientes porque si existen constantes $k_1, k_2, \cdots, k_n$ tales que $k_1e_1 + k_2e_2 + \cdots + k_ne_n = 0$, entonces $(k_1, k_2, \cdots, k_n) = 0$ y así $k_1 = k_2 = ... = k_n = 0$. Además, cualquier vector $x = (x_1, x_2, \cdots, x_n)$ puede ser generado por los vectores $e_1, e_2, \cdots, e_n$, pues $(x_1, x_2, \cdots, x_n) = x_1e_1 + x_2e_2 + \cdots + x_ne_n$.

Ejemplo 33.

a) El conjunto $\beta = \{(1, 0), (0, 1)\}$ es la base canónica de $\mathbb{R}^2$.

b) El conjunto $\beta = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ corresponde a la base canónica para $\mathbb{R}^3$.

c) ¿Cuáles son, explícitamente, los cuatro vectores de la base canónica para $\mathbb{R}^4$?

Nota 5.
Una mirada geométrica al plano $\mathbb{R}^2$ o al espacio $\mathbb{R}^3$, bastaría para convencernos de que las bases no pueden ser únicas. Pueden existir un número infinito de ellas y esto lo haremos explícito enseguida. Además, tal vez pueda ser claro para el lector que cada elemento de la base genera uno de los *ejes* respecto a los cuales se describirán *todos* los demás vectores del plano o del espacio. Por ejemplo, el típico plano $\mathbb{R}^2$ con ejes ortogonales es sólo una forma de describir los demás puntos del plano: la base típica es la canónica $\beta = \{(1, 0), (0, 1)\}$

Ejemplo 34. (Las bases no son únicas)
Además de la base canónica, veamos que el conjunto $\beta = \{(1, 1, 0), (0, 1, 1), (1, 1, 1)\}$ *también* es una base para $\mathbb{R}^3$.

Solución

Sea $(x, y, z) \in \mathbb{R}^3$. Entonces una aplicación del método gaussiano nos muestra que (x, y, z) se puede expresar como combinación lineal de $(1,1,0)$, $(0,1,1)$ y $(1,1,1)$:

$$(x, y, z) = (y - z)(1,1,0) + (y - x)(0,1,1) + (x - y + z)(1,1,1)$$

Además, los vectores de β son linealmente independientes porque si para algunos k_1, k_2, k_3 tuviéramos

$$k_1(1,1,0) + k_2(0,1,1) + k_3(1,1,1) = (0,0,0)$$

entonces también tendríamos que k_1, k_2, k_3 satisfacen el sistema de ecuaciones lineales

$$\begin{aligned} k_1 + k_3 &= 0 \\ k_1 + k_2 + k_3 &= 0 \\ k_2 + k_3 &= 0 \end{aligned}$$

y llegaríamos a que $k_1 = k_2 = k_3 = 0$, que a su vez implica que el conjunto

$$\beta = \{(1,1,0), (0,1,1), (1,1,1)\}$$

forma una base para $\mathbb{R}^3$. Es claro, entonces, que $\mathbb{R}^3$ (y, en general, $\mathbb{R}^n$) tiene *infinitas bases.*

Ejemplo 35.
El conjunto $\beta = \{1, x, x^2, \ldots, x^n\}$ es una base para el espacio P_n de los polinomios de grado menor o igual que n.

Solución

a) β es linealmente independiente, pues si para ciertos escalares $a_0, a_1, ..., a_n \in \mathbb{R}$ se tiene que

$$a_0 + a_1 x + \cdots + a_n x^n \equiv 0$$

para todo x, entonces $a_0 = a_1 = \cdots = a_n = 0$ (¿Podría el lector llenar los detalles que nos conducen a esta conclusión? Recuerde cuándo dos polinomios son iguales).

b) $\langle \beta \rangle = P_n$, pues, de hecho, todo polinomio de grado $\leq n$ se puede escribir como

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n$$

para algunos $a_0, a_1, \cdots, a_n \in \mathbb{R}$. ▲

Buscando describir las características esenciales de una base, tenemos el siguiente teorema:

Teorema 3.
Sea V un espacio vectorial tal que $\beta = \{\beta_1, \ldots, \beta_m\}$ es una base para algunos $\beta_1, \beta_2, \ldots, \beta_m \in V$. Entonces cualquier conjunto de vectores $\alpha = \{\alpha_1, \ldots, \alpha_n\}$ linealmente independientes en V no puede contener más de m elementos; es decir, $n \leq m$.

Demostración

a) Consideremos el conjunto $\{\alpha_1, \beta_1, \ldots, \beta_m\}$. Este es, claramente, linealmente dependiente, pues α_1 es combinación lineal de los β_i's. Pero en tal caso, también algún β_j es combinación de los vectores anteriores en la lista del conjunto de arriba. Eliminemos este β_j.

b) Ahora consideremos el conjunto $\{\alpha_1, \alpha_2, \beta_1, \ldots, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_m\}$. Éste es también linealmente dependiente y, por tanto, algún β_i es combinación de los vectores anteriores en el conjunto inmediatamente anterior. Eliminemos también a β_i.

c) Procediendo de esta forma podemos ir *reemplazando* los vectores de β por vectores de α. Si sucediera que pudiéramos reemplazar *todos* los β's por algunos (no todos) de los α's, entonces tendríamos que el conjunto α sería linealmente dependiente en V, y esta es una contradicción. ■

Corolario 1 (*Todas las bases tienen el mismo número de vectores*)
Si V es un espacio vectorial con una base de m elementos, entonces cualquier otra base también tiene m elementos.

Demostración
Si β es una base de V con m elementos y α es otra con n elementos, entonces, puesto que los elementos de β son linealmente independientes, se tiene (por el teorema 3) que $n \geq m$. De manera similar, puesto que los elementos de α son linealmente independientes, entonces $m \geq n$; de estas dos desigualdades se tiene que $m = n$. ■

Puesto que todas las bases tienen el mismo número de elementos, el anterior corolario nos permite entonces definir el concepto central de *dimensión* de un espacio vectorial.

Definición 7. (Dimensión de un espacio vectorial)
Si V es un espacio vectorial con al menos una base de m elementos, entonces diremos que m es la *dimensión* de V, y escribiremos

$$\dim V = m$$

Si $V = \{ 0 \}$, entonces diremos que $\dim V = 0$.

Nota 6.
Si $m = 0, 1, 2, \ldots$ es la dimensión de V diremos que V es *finito-dimensional.* En otro caso, diremos que V es *infinito-dimensional* o, simplemente, de *dimensión infinita.*

Ejemplo 36.

a) El espacio vectorial $\mathbb{R}^n$ tiene dimensión n. Una base para $\mathbb{R}^n$ la conforman los vectores unitarios $e_1 = (1, 0, ..., 0)$, $e_2 = (0, 1, ..., 0)$, ... , $e_n = (0, 0, ..., 1)$.

b) El espacio vectorial $\mathfrak{M}_{m\times n}$ tiene dimensión mn. Una base para el conjunto $\mathfrak{M}_{m\times n}$ es

$$\left\{ \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \right\}$$

c) El espacio vectorial P_n tiene dimensión $n + 1$. ¿Cuál es una base para este espacio?

d) El espacio vectorial $\{ k(x_0, y_0) / k \in \mathbb{R}) \}$ con $(x_0, y_0) \neq (0, 0)$ tiene dimensión 1; una base es $\{(x_0, y_0)\}$.

e) El espacio vectorial $\{ k(x_0, y_0, z_0) + l(x_1, y_1, z_1) / k, l \in \mathbb{R}) \}$ en $\mathbb{R}^3$ tiene dimensión 2 si (x_0, y_0, z_0) y (x_1, y_1, z_1) son linealmente independientes; tiene dimensión 1 si (x_0, y_0, z_0) y (x_1, y_1, z_1) son linealmente dependientes y uno de los dos es diferente de $(0, 0, 0)$; y tiene dimensión 0 si ambos vectores son $(0, 0, 0)$.

f) Se puede mostrar (y lo haremos en el ejemplo 41) que el espacio de todas las funciones $f : S \longrightarrow \mathbb{R}$, $S \subseteq \mathbb{R}$, $S \neq \phi$, es *infinito-dimensional*; y que también lo es el espacio de *todos* los polinomios (sin grado específico).

Corolario 2 (*Una base "minimiza" la información lineal*)
Si dim $V = m$, entonces

a) *Cualquier subconjunto de V con más de m vectores es linealmente dependiente.*

b) *Ningún subconjunto de V con menos de m vectores puede generar a V.*

Demostración

La parte a) se desprende del teorema 3. Para la parte b) supongamos que, efectivamente, un subconjunto con menos de m vectores genera a V. Entonces, de allí podemos extraer un subconjunto linealmente independiente que también genera a V; en efecto: ordenemos los vectores del conjunto de la forma $\{\beta_1, ..., \beta_k\}$ y comencemos a eliminar, de derecha a izquierda, todos aquellos vectores que son combinaciones lineales de los vectores que han quedado en el conjunto luego de la anterior eliminación. Al final debemos obtener un conjunto linealmente independiente, y que además genera a V, lo que es una contradicción, pues esta sería entonces una base para V con menos de m elementos. ■

Ejemplo 37.

Un conjunto de cuatro vectores en $\mathbb{R}^3$ tal como

$$S = \{ (1, 2, -3), (2, 1, -3), (2, -3, 4), (4, 7, -6) \},$$

es, de acuerdo con el corolario 2, linealmente dependiente, pues $\dim \mathbb{R}^3 = 3$. De hecho,

$$c_1(1, 2, -3) + c_2(2, 1, -3) + c_3(2, -3, 4) + c_4(4, 7, -6) = 0$$

para $c_1 = -\frac{70}{9}$, $c_2 = \frac{32}{9}$, $c_3 = -\frac{5}{3}$, $c_4 = 1$.

Ejemplo 38.

Confirmando el corolario 2, el conjunto de vectores $\{ (1, 0, 0), (0, 1, 0) \}$ no genera $\mathbb{R}^3$ pues, por ejemplo, el vector $(3, 4, 2)$ no puede ser expresado como una combinación lineal de los vectores $(1, 0, 0)$ y $(0, 1, 0)$ como puede comprobarse fácilmente.

Teorema 4.

a) *Cualquier subconjunto linealmente independiente de un espacio finito-dimensional V puede completarse para hacer de él una base.*

b) *Si $\dim V = m$, entonces una colección de m vectores en V linealmente independientes forma una base para V.*

c) *Si $\dim V = m$, entonces una colección de m vectores que genere a V forma una base para V.*

Demostración

Probaremos la parte a) y la b). La parte c) queda como ejercicio para el lector.

a) Sea β un conjunto linealmente independiente de V. Si $\langle \beta \rangle = V$ hemos terminado. Si no, sea $x \in V - \langle \beta \rangle$. Entonces pueden suceder dos situaciones: $\langle \beta \cup \{ x \} \rangle = V$ y en tal caso habríamos finalizado la prueba. Si no, repetimos el proceso inicial de adicionar vectores. Que este proceso no puede seguir indefinidamente es resultado de que V es finito-dimensional, y de que cualquier conjunto de más elementos que los de una base es linealmente dependiente (Corolario 2a)).

b) Si no fuera una base, entonces, utilizando la parte a) de este teorema, podríamos completarla para hacer de ella una base. Pero esto contradiría el hecho de que todas las bases tienen el mismo número de elementos. $\blacksquare$

Ejemplo 39.
Hallemos una base de $\mathbb{R}^2$ que contenga al vector $(3, -5)$.

Solución

De acuerdo con el teorema anterior, es suficiente encontrar un vector (x, y) que sea linealmente independiente del vector $(3, -5)$; es decir, no debe existir $k \in \mathbb{R}$, $k \neq 0$, tal que

$$
\begin{aligned}
x &= 3k \\
y &= -5k
\end{aligned}
$$

Tomemos, por ejemplo, $x = 3$, $y = 5$. Observemos que no existe $k \in \mathbb{R}$ tal que $3 = 3k$ y $5 = -5k$. Luego, el conjunto $\{ (3, -5), (3, 5) \}$ forma una base para $\mathbb{R}^2$.

Ejemplo 40.
Hallemos una base de $\mathbb{R}^3$ que contenga a los vectores $(2, 1, 3)$ y $(3, 1, -5)$.

Solución

De acuerdo con la parte b) del teorema 4 es suficiente encontrar un vector (x, y, z) que sea linealmente independiente de los vectores $(2, 1, 3)$, $(3, 1, -5)$. Por tanto, si existen escalares k_1, k_2, k_3 tales que $k_1(2, 1, 3) + k_2(3, 1, -5) + k_3(x, y, z) = 0$, entonces $k_1 = k_2 = k_3 = 0$. Podemos reescribir la igualdad anterior como el siguiente sistema de ecuaciones lineales homogéneo:

$$
\begin{aligned}
2k_1 + 3k_2 + xk_3 &= 0 \\
k_1 + k_2 + yk_3 &= 0 \\
3k_1 - 5k_2 + zk_3 &= 0
\end{aligned}
$$

Veamos qué condición debe satisfacer el vector (x, y, z) para que la única solución de este sistema de ecuaciones lineales sea precisamente $k_1 = k_2 = k_3 = 0$. La matriz aumentada de este sistema es

$$\begin{array}{c} \left[\begin{array}{rrc|c} 2 & 3 & x & 0 \\ 1 & 1 & y & 0 \\ 3 & -5 & z & 0 \end{array}\right] \end{array} \begin{array}{l} F_1 \\ F_2 \\ F_3 \end{array}$$

Utilizando de nuevo el método de eliminación gaussiana obtenemos que

$$\left[\begin{array}{ccc|c} 1 & \frac{3}{2} & \frac{x}{2} & 0 \\ 1 & 1 & y & 0 \\ 0 & -8 & -3y+z & 0 \end{array}\right] \begin{array}{l} F_1 \longleftrightarrow \frac{1}{2}F_1 \\ \\ F_3 \longleftrightarrow F_3 - 3F_2 \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & \frac{3}{2} & \frac{x}{2} & 0 \\ 0 & \frac{1}{2} & \frac{x}{2} - y & 0 \\ 0 & -8 & -3y+z & 0 \end{array}\right] \quad F_2 \longleftrightarrow F_1 - F_2$$

$$\left[\begin{array}{ccc|c} 1 & \frac{3}{2} & \frac{x}{2} & 0 \\ 0 & 1 & x - 2y & 0 \\ 0 & -8 & -3y+z & 0 \end{array}\right] \quad F_2 \longleftrightarrow 2F_2$$

$$\left[\begin{array}{ccc|c} 1 & 0 & -x+3y & 0 \\ 0 & 1 & x-2y & 0 \\ 0 & 0 & 8x-19y+z & 0 \end{array}\right] \begin{array}{l} F_1 \longleftrightarrow F_1 - \frac{3}{2}F_2 \\ \\ F_3 \longleftrightarrow F_3 + 8F_2 \end{array}$$

Por tanto, para que (x, y, z) sea linealmente independiente de los vectores $(2, 1, 3)$ y $(3, 1, -5)$ es necesario que $8x - 19y + z \neq 0$. Tomemos, por ejemplo, el vector $(1, 0, -1)$, el cual satisface la condición requerida. Así, el conjunto $\{(2, 1, 3), (3, 1, -5), (1, 0, -1)\}$ forma una base para $\mathbb{R}^3$. Claramente, la ecuación $8x - 19y + z = 0$ describe, en $\mathbb{R}^3$, el plano generado por los vectores $(2, 1, 3)$ y $(3, 1, -5)$; de hecho, el vector normal al plano, que es $(8, -19, 1)$, es ortogonal a $(2, 1, 3)$ y $(3, 1, -5)$.

Teorema 5. ***(La dimensión respeta el orden)***
Si V es finito-dimensional y W es un subespacio de V, entonces W es finito-dimensional y, además,

$$\dim W \leq \dim V$$

Demostración

Sea V un espacio de dimensión m. Si β es un conjunto de vectores linealmente independientes en W, entonces β también es un conjunto de vectores linealmente independientes en V. Por la parte a) del corolario 2, todo conjunto de vectores linealmente independientes en W tiene a lo más m elementos. Por tanto, $\dim W \leq m$. ■

Corolario 3

Si W es subespacio de V, y W es infinito-dimensional, entonces también V es infinito-dimensional.

Ejemplo 41. (Peano (1888))
El conjunto de *todos* los polinomios es *infinito-dimensional* pues si $\beta = \{ p_1, \ldots, p_n \}$ es una base, entonces el monomio $q(x) = x^{m+1}$, donde m es el máximo grado de los polinomios en β *no* puede expresarse como combinación lineal de los polinomios en β. Ahora: como el conjunto de todos los polinomios está incluido en el de todas las funciones $f : \mathbb{R} \longrightarrow \mathbb{R}$, entonces (por el corolario 3) este espacio también es *infinito-dimensional.*

Nota 7.
Como ya puede haberse apreciado de los ejemplos en esta lección, es todavía posible que un espacio infinito-dimensional tenga subespacios finito-dimensionales. ¿Podría el lector señalar un ejemplo concreto de esto?

Ejercicios 2

1) Determine si los vectores $(4, 3, 1, -1)$, $(3, 0, 1, -2)$, $(3, 2, 1, 2)$, y $(5, 0, 2, -2)$ son linealmente independientes.

2) ¿Para qué valor de α son linealmente dependientes los vectores $(2, 4, 6)$, $(2, -1, 4)$ y $(2, 1, \alpha)$?

3) ¿Será que el conjunto $\{ (1, 1, 0), (2, 3, -1), (5, 1, 1) \}$ es un conjunto linealmente independiente en $\mathbb{R}^3$? ¿Será una base para $\mathbb{R}^3$?

4) Encuentre dos bases diferentes de $\mathbb{R}^3$, pero con *dos* de sus tres vectores iguales.

5) De los seis vectores $(0, 0, 0)$, $(0, 1, 0)$, $(0, 2, 0)$, $(0, 0, 1)$, $(0, 2, 1)$, $(0, 0, 1)$, ¿cuáles grupos de a tres vectores conforman bases para $\mathbb{R}^3$?

6) Construya una base para el subespacio de $\mathbb{R}^3$

$$W = \{ (x, y, z) \in \mathbb{R}^3 \,/\, 3x + y - z = 0 \}$$

3. Bases ortonormales para $\mathbb{R}^n$

El concepto de ortogonalidad en vectores, ya vimos en la lección anterior, está enraizado profundamente en la geometría de los antiguos griegos. Si la idea de base es la de simplificar la "información lineal" que poseemos del espacio vectorial bajo estudio, la de ortogonalidad es la de hacer aún más clara la conexión de la geometría de un espacio vectorial con su álgebra. Es decir, se busca una base en la cual cada construcción *geométrica* se pueda convertir en una *algebraica* más simple y así reducir los cálculos notablemente.

Definición 8. (Conjunto de vectores ortogonales)

a) Diremos que $S = \{\beta_1, \ldots, \beta_m\}$ es un *conjunto de vectores ortogonales* en $\mathbb{R}^n$ si $\beta_i \cdot \beta_j = 0$ para todo $i \neq j$, $1 \leq i, j \leq m$.

b) Y diremos que el conjunto de vectores ortogonales S es *ortonormal* si $||\,\beta_i\,|| = 1$ para todo $1 \leq i \leq m$.

Ejemplo 42.

a) El conjunto de vectores $\left\{ \left(\frac{4}{5}, \frac{3}{5}\right), \left(-\frac{3}{5}, \frac{4}{5}\right) \right\}$ es un conjunto de vectores ortonormales en $\mathbb{R}^2$, como el lector fácilmente puede comprobar.

b) El conjunto de vectores $\{(1,0,0), (0,1,0), (0,0,1)\}$ es un conjunto de vectores ortonormales en $\mathbb{R}^3$.

c) El conjunto de vectores $\left\{ \left(\frac{2}{3}, \frac{1}{3}, -\frac{2}{3}\right), \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right), \left(\frac{2}{3}, -\frac{2}{3}, \frac{1}{3}\right) \right\}$ es un conjunto de vectores ortonormales en $\mathbb{R}^3$.

d) El conjunto de vectores $\{(1,4,5), (-3,2,-1), (-1,-1,1)\}$ es un conjunto de vectores ortogonales en $\mathbb{R}^3$. Sin embargo, estos vectores no son ortonormales porque, por ejemplo, $||\,(1,4,5)\,|| = \sqrt{42}$.

Teorema 6.
Todo conjunto $S = \{\beta_1, \ldots, \beta_n\}$ de vectores ortogonales no-nulos en $\mathbb{R}^n$ es una base para $\mathbb{R}^n$.

Demostración
Por la parte b) del teorema 4, es suficiente probar que S es linealmente independiente. Para ello, supongamos que existen $k_1, k_2, \ldots, k_n \in \mathbb{R}$ tales que

$$k_1\beta_1 + k_2\beta_2 + \cdots + k_n\beta_n = 0$$

Multiplicando a ambos lados de esta igualdad (con el producto interior) por β_i $(1 \leq i \leq n)$, se obtiene

$$k_1\beta_1 \cdot \beta_i + k_2\beta_2 \cdot \beta_i + \cdots + k_i\beta_i \cdot \beta_i + \cdots + k_n\beta_n \cdot \beta_i = 0$$

Y como $\beta_j \cdot \beta_i = 0$ si $j \neq i$, entonces se tiene que

$$k_i \|\, \beta_i \,\|^2 = 0;$$

pero puesto que $\beta_i \neq 0$, entonces $k_i = 0$; y esto demuestra la independencia lineal de S. ■

Definición 9. (Base ortogonal)
Una *base ortogonal* para $\mathbb{R}^n$ es un conjunto de n vectores ortogonales no-nulos. Si, además, cada uno de estos vectores tiene norma 1, entonces al conjunto se le llama una *base ortonormal* para $\mathbb{R}^n$.

Nota 8. (Proyección ortogonal)
Observemos que si $S = \{\beta_1, \ldots, \beta_n\}$ es una base ortonormal para $\mathbb{R}^n$, y $\beta = k_1\beta_1 + \cdots + k_n\beta_n$ con $k_i \in \mathbb{R}$ $i = 1, 2, \ldots$, entonces $\beta \cdot \beta_i = k_i \|\, \beta_i \,\|^2 = k_i$; es decir, los coeficientes k_i son el producto interior de β y β_i.

El siguiente es un proceso estándar para construir una base ortonormal en $\mathbb{R}^n$ conocido como el *proceso de ortogonalización de Gram-Schmidt*:

Teorema 7. ***(Proceso de Gram-Schmidt (Schmidt (1907)))***
Todo subespacio no-nulo de $\mathbb{R}^n$ tiene una base ortonormal. De hecho, si $S = \{\beta_1, \ldots, \beta_m\}$ es una base para un subespacio dado de $\mathbb{R}^n$, donde $0 \leq m \leq n$, entonces el conjunto de vectores $\{\alpha_1, \ldots, \alpha_m\}$ definido mediante recursión así:

$$\begin{aligned}
\alpha_1 &= \beta_1 \\
\alpha_2 &= \beta_2 - \frac{\alpha_1 \cdot \beta_2}{\|\, \alpha_1 \,\|^2}\,\alpha_1 \\
\alpha_3 &= \beta_3 - \frac{\alpha_1 \cdot \beta_3}{\|\, \alpha_1 \,\|^2}\,\alpha_1 - \frac{\alpha_2 \cdot \beta_3}{\|\, \alpha_2 \,\|^2}\,\alpha_2 \\
\alpha_4 &= \beta_4 - \frac{\alpha_1 \cdot \beta_4}{\|\, \alpha_1 \,\|^2}\,\alpha_1 - \frac{\alpha_2 \cdot \beta_4}{\|\, \alpha_2 \,\|^2}\,\alpha_2 - \frac{\alpha_3 \cdot \beta_4}{\|\, \alpha_3 \,\|^2}\,\alpha_3 \\
\vdots\ & \qquad \vdots \\
\alpha_m &= \beta_m - \sum_{i=1}^{m-1} \frac{\alpha_i \cdot \beta_m}{\|\, \alpha_i \,\|^2}\,\alpha_i
\end{aligned}$$

es un conjunto ortogonal que, después de normalizarse, se convierte en una base ortonormal del subespacio.

Demostración

Primero se elige $\alpha_1 = \beta_1$. Después buscamos α_2 en el espacio generado por β_1 y β_2; es decir, por α_1 y β_2. Encontremos entonces $k_1,\ k_2$ tales que $\alpha_2 = k_1\beta_2 + k_2\alpha_1$ y $\alpha_1 \cdot \alpha_2 = 0$. Aquí, $\alpha_2 \cdot \alpha_1 = k_1\beta_2 \cdot \alpha_1 + k_2\alpha_1 \cdot \alpha_1$, y así $k_2 = -\dfrac{\beta_2 \cdot \alpha_1}{||\,\alpha_1\,||^2}\, k_1$; y tomando $k_1 = 1$ (¿por qué podemos hacerlo?) obtenemos que $k_2 = -\dfrac{\beta_2 \cdot \alpha_1}{||\,\alpha_1\,||^2}$ y así

$$\alpha_2 = \beta_2 - \frac{\alpha_1 \cdot \beta_2}{||\,\alpha_1\,||^2}\, \alpha_1$$

Luego buscamos un α_3 en el espacio generado por $\beta_1, \beta_2, \beta_3$ que sea ortogonal a $\alpha_1,\ \alpha_2$; es decir, generado por $\alpha_1,\ \alpha_2$ y β_3 y tal que $\alpha_3 \cdot \alpha_1 = \alpha_3 \cdot \alpha_2 = 0$. Nuevamente escribimos

$$\alpha_3 = k_1\beta_3 + k_2\alpha_1 + k_3\alpha_2 \tag{1}$$

Podemos, como antes, asumir $k_1 = 1$. Entonces, multiplicando interiormente por α_1 y α_2 respectivamente, obtenemos que

$$\begin{aligned} \alpha_3 \cdot \alpha_1 &= \beta_3 \cdot \alpha_1 + k_2\alpha_1 \cdot \alpha_1 + k_3\alpha_2 \cdot \alpha_1 \\ \alpha_3 \cdot \alpha_2 &= \beta_3 \cdot \alpha_2 + k_2\alpha_1 \cdot \alpha_2 + k_3\alpha_2 \cdot \alpha_2 \end{aligned}$$

o, lo que es igual,

$$\begin{aligned} 0 &= \beta_3^{\cdot}\alpha_1 + k_2||\,\alpha_1\,||^2 \\ 0 &= \beta_3^{\cdot}\alpha_2 + k_3||\,\alpha_2\,||^2 \end{aligned}$$

Luego $k_2 = -\dfrac{\alpha_1 \cdot \beta_3}{||\,\alpha_1\,||^2}$ y $k_3 = -\dfrac{\alpha_2 \cdot \beta_3}{||\,\alpha_2\,||^2}$. Por tanto,

$$\alpha_3 = \beta_3 - \frac{\alpha_1 \cdot \beta_3}{||\,\alpha_1\,||^2}\,\alpha_1 - \frac{\alpha_2 \cdot \beta_3}{||\,\alpha_2\,||^2}\,\alpha_2$$

El proceso continuará de esta manera hasta obtener una base ortogonal $\{\,\alpha_i\,\}$ y, de allí, por normalización, una base ortonormal. ■

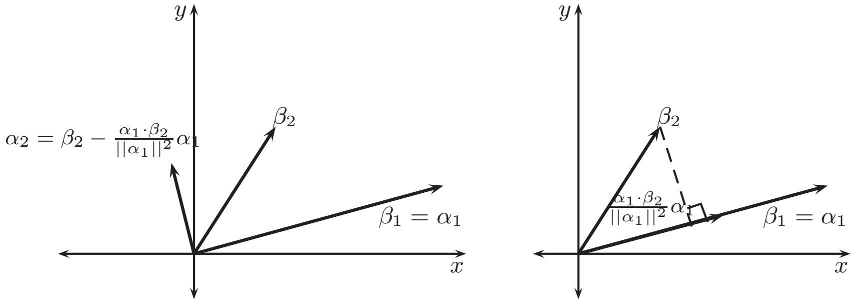


Figura 6. Proceso de Gram-Schmidt

Ejemplo 43.
Sean $\beta_1 = (1, 1, 0)$, $\beta_2 = (1, 0, 1)$ y $\beta_3 = (0, 1, 1)$. Construyamos una base ortonormal para $\mathbb{R}^3$ a partir de estos tres vectores.

Solución
Sea $\alpha_1 = (1, 1, 0)$; puesto que $\dfrac{\alpha_1 \cdot \beta_2}{||\,\alpha_1\,||^2} = \dfrac{1}{2}$, entonces

$$\alpha_2 = (1, 0, 1) - \frac{1}{2}(1, 1, 0) = \left(\frac{1}{2}, -\frac{1}{2}, 1\right)$$

El tercer eje perpendicular surge de calcular $\dfrac{\alpha_1 \cdot \beta_3}{||\,\alpha_1\,||^2} = \dfrac{1}{2}$, $\dfrac{\alpha_2 \cdot \beta_3}{||\,\alpha_2\,||^2} = \dfrac{1}{3}$, pues así

$$\alpha_3 = (0, 1, 1) - \frac{1}{2}(1, 1, 0) - \frac{1}{3}\left(\frac{1}{2}, -\frac{1}{2}, 1\right) = \left(-\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right)$$

La base ortonormal estará entonces conformada por

$$\alpha_1^* = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right),\ \alpha_2^* = \left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \sqrt{\frac{2}{3}}\right),\ \alpha_3^* = \left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$$

Si, por ejemplo, consideramos el vector $(7, 1, 2)$, entonces $k_1\alpha_1^* + k_2\alpha_2^* + k_3\alpha_3^* = (7, 1, 2)$, donde

$$k_1 = (7, 1, 2) \cdot \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right) = \frac{8}{\sqrt{2}}$$
$$k_2 = (7, 1, 2) \cdot \left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \sqrt{\frac{2}{3}}\right) = \frac{10}{\sqrt{6}}$$
$$k_3 = (7, 1, 2) \cdot \left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) = -\frac{4}{\sqrt{3}}$$

Ejemplo 44.
Construyamos una base ortonormal de $\mathbb{R}^3$ a partir de la base $\beta = \{\,\beta_1, \beta_2, \beta_3\,\}$, donde

$$\beta_1 = (\,2, 1, 2\,),\ \ \beta_2 = (\,2, 1, 1\,),\ \ \beta_3 = (\,0, 1, 1\,)$$

Solución
Sea $\alpha_1 = \beta_1 = (\,2, 1, 2\,)$. Calculemos ahora α_2 y α_3:

$$\alpha_2 = \beta_2 - \left(\frac{\alpha_1 \cdot \beta_2}{||\,\alpha_1\,||^2} \right) \alpha_1 = (\,2, 1, 1\,) - \frac{7}{9}(\,2, 1, 2\,) = \left(\frac{4}{9}, \frac{2}{9}, -\frac{5}{9} \right)$$

$$\begin{aligned}
\alpha_3 &= (\,0, 1, 1\,) - \left(\frac{\alpha_1 \cdot \beta_3}{||\,\alpha_1\,||^2} \right) \alpha_1 - \left(\frac{\alpha_2 \cdot \beta_3}{||\,\alpha_2\,||^2} \right) \alpha_2 \\
&= (\,0, 1, 1\,) - \frac{3}{9}(\,2, 1, 2\,) - \left(-\frac{\frac{3}{9}}{\frac{5}{9}} \right) \left(\frac{4}{9}, \frac{2}{9}, -\frac{5}{9} \right) = \left(-\frac{2}{5}, \frac{4}{5}, 0 \right)
\end{aligned}$$

Por tanto,

$$\alpha_1^* = \left(\frac{2}{3}, \frac{1}{3}, \frac{2}{3} \right), \quad \alpha_2^* = \left(\frac{4}{\sqrt{45}}, \frac{2}{\sqrt{45}}, -\frac{5}{\sqrt{45}} \right),$$

$$\alpha_3^* = \left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, 0 \right)$$

Nota 9. (Sobre el origen del proceso Gram-Schmidt)
Aunque el método Gram-Schmidt es usualmente asociado con Jorgen P. Gram (1883) y Erhard Schmidt (1907), ellos no fueron, sin embargo, los primeros en utilizar este método. El proceso parece ser un resultado de Laplace (1812), que también fue utilizado de manera recurrente por Cauchy en 1836.

Finalmente, y para cerrar esta sección, construimos un tipo particular de matrices que están conformadas en sus filas por bases ortonormales. Este tipo de matrices tiene numerosas aplicaciones en geometría y física, y sobre ellas discutiremos en las próximas lecciones.

Definición 10. (Matriz ortogonal)
Una matriz cuadrada A de tamaño $n \times n$ es *ortogonal* si

$$A^T A = I_n = AA^T \quad \text{y, por tanto,} \quad A^{-1} = A^T$$

Teorema 8. ***(Sobre matrices ortogonales)***

a) *Una matriz A de tamaño $n \times n$ es ortogonal si, y sólo si, sus vectores columna (y también sus vectores fila) forman una base ortonormal para $\mathbb{R}^n$.*

b) *El determinante de una matriz ortogonal A tiene valor 1 ó -1.*

Demostración

a) Esta es una conclusión directa de la condición de ortogonalidad $A^TA = AA^T = I_n$.

b) Como $A^TA = I_n$ y $\det A^T = \det A$ entonces $(\det A)^2 = 1$. Así, $\det A = 1$ ó $\det A = -1$. ■

Ejemplo 45.

Veamos si la matriz $A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{\sqrt{2}}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}$ es ortogonal.

Solución

Tenemos que $A^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}$; por tanto, $A^TA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = AA^T$, lo que muestra la ortogonalidad de la matriz A.

Ejemplo 46.

Veamos si la matriz $A = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{4}{\sqrt{45}} & \frac{2}{\sqrt{45}} & -\frac{5}{\sqrt{45}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \end{bmatrix}$ es ortogonal.

Solución

Tenemos que $A^T = \begin{bmatrix} \frac{2}{3} & \frac{4}{\sqrt{45}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{3} & \frac{2}{\sqrt{45}} & \frac{2}{\sqrt{5}} \\ \frac{2}{3} & -\frac{5}{\sqrt{45}} & 0 \end{bmatrix}$; por tanto, $A^TA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = AA^T$, lo que demuestra que A es, efectivamente, una matriz ortogonal.

Ejercicios 3

1) Muestre que la base canónica es ortonormal en $\mathbb{R}^n$.

2) Muestre que el conjunto $\left\{ \left(\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}} \right), \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0 \right), \left(\frac{1}{\sqrt{30}}, -\frac{2}{\sqrt{30}}, -\frac{5}{\sqrt{30}} \right) \right\}$ es una base ortonormal para $\mathbb{R}^3$.

3) A partir del vector $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$ complete una base ortonormal para $\mathbb{R}^2$.

4) Pruebe que a partir de los vectores $(1,1,0)$, $(1,0,1)$ y $(0,1,1)$ se puede construir una base ortonormal de $\mathbb{R}^3$ conformada por los vectores $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \right)$, $\left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \sqrt{\frac{2}{3}} \right)$ y $\left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right)$.

5) A partir de los vectores $(1,2,1)$, $(1,-1,0)$, y $(3,2,4)$ calcule una base ortonormal para $\mathbb{R}^3$.

6) Calcule una base ortonormal para los siguientes planos en $\mathbb{R}^3$:

   a) $\{ (x,y,z) \in \mathbb{R}^3 \,/\, 3x + 2y + z = 0 \}$
   b) $\{ (x,y,z) \in \mathbb{R}^3 \,/\, x - y + z = 0 \}$

7) Decida si las siguientes matrices son (o no) ortogonales:

$$\text{a)} \quad \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix} \qquad \text{b)} \quad \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

$$\text{c)} \quad \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix} \quad \text{donde } \ 0 \le \theta \le \pi.$$

*8) ¿Cómo podría el lector relacionar el concepto de proyección (ortogonal) de un vector sobre otro (lección 4) con la noción de proyección ortogonal desarrollada en la Nota 8 anterior?

4. Bases para el espacio-solución de un sistema de ecuaciones lineales homogéneo

Ya sabemos que el sistema lineal homogéneo

$$AX = 0, \quad A = [\, a_{ij} \,]_{m\times n}, \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

induce la pregunta sobre el comportamiento del subespacio vectorial de sus soluciones en $\mathbb{R}^n$. ¿Qué luces nos puede aportar lo desarrollado hasta ahora en esta lección con respecto a este espacio de soluciones? Este espacio es de dimensión menor o igual que n (la dimensión de $\mathbb{R}^n$) y además podemos describir una base para él. Ilustremos esto en los siguientes ejemplos.

Ejemplo 47.
El sistema

$$\begin{aligned} x + 2y - 3z &= 0 \\ 2x + y - 3z &= 0 \end{aligned}$$

tiene soluciones (después de aplicar el método gaussiano) de la forma

$$\begin{aligned} x &= k \\ y &= k \\ z &= k \end{aligned} \qquad \text{o} \qquad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = k \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad k \in \mathbb{R}$$

Así, el espacio solución es $\langle\, (1, 1, 1)\, \rangle$ y tiene dimensión 1: es la recta

$$\frac{x}{1} = \frac{y}{1} = \frac{z}{1}$$

Ejemplo 48.
Mediante el método gaussiano, el lector puede mostrar que el sistema

$$\begin{aligned} x_1 + 2x_2 - 3x_4 + x_5 &= 0 \\ x_1 + 2x_2 + x_3 - 3x_4 + x_5 + 2x_6 &= 0 \\ x_1 + 2x_2 - 3x_4 + 2x_5 + x_6 &= 0 \\ 3x_1 + 6x_2 + x_3 - 9x_4 + 4x_5 + 3x_6 &= 0 \end{aligned}$$

tiene como soluciones

$$\begin{aligned} x_1 &= r + 3s - 2t \\ x_2 &= t \\ x_3 &= -2r \\ x_4 &= s \\ x_5 &= -r \\ x_6 &= r \end{aligned}$$

para r, s, $t \in \mathbb{R}$. Es decir,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = r \begin{bmatrix} 1 \\ 0 \\ -2 \\ 0 \\ -1 \\ 1 \end{bmatrix} + s \begin{bmatrix} 3 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

y así, el espacio-solución del anterior sistema lineal homogéneo es el espacio vectorial generado $\langle\, (1, 0, -2, 0, -1, 1), \ (3, 0, 0, 1, 0, 0), \ (-2, 1, 0, 0, 0, 0)\, \rangle$ que, claramente, tiene dimensión 3: es un hiperplano (de dimensión 3) en $\mathbb{R}^6$.

Ejemplo 49.
Consideremos el sistema lineal

$$\begin{aligned} 2x + 3y + 5z &= 0 \\ -x + 7y - z &= 0 \\ 4x - 11y + z &= 0 \end{aligned}$$

Aplicando el método gaussiano se puede ver fácilmente que la única solución de este sistema lineal homogéneo es $(0, 0, 0)$. Por tanto, el espacio-solución de este sistema es de dimensión cero: $(0, 0, 0)$ es la intersección de los tres planos de arriba.

Ejemplo 50.
El sistema lineal homogéneo

$$\begin{aligned} 2x + 3y &= 0 \\ x + y &= 0 \end{aligned}$$

tiene como única solución a $(0, 0)$. Por tanto, el espacio-solución de este sistema lineal también es de dimensión cero: $(0, 0)$ es la intersección de las dos rectas.

Ejemplo 51.
Consideremos el sistema lineal

$$\begin{aligned} -x + 2y &= 0 \\ 2x - 4y &= 0 \end{aligned}$$

Aplicando el método gaussiano se tiene que las soluciones de este sistema son de la forma $x = k$, $y = k/2$ para todo $k \in \mathbb{R}$. El espacio-solución de este sistema lineal es $\langle\, (1, \frac{1}{2})\, \rangle$ y su dimensión es 1: es la recta $y = x/2$. ▲

Podemos entonces concluir la presente sección resumiendo la información que, hasta ahora, tenemos sobre los sistemas lineales homogéneos.

Teorema 9. ***(¿Qué sabemos, hasta ahora, de los sistemas homogéneos?)***
Sea A una matriz $n \times n$. Entonces las cinco afirmaciones siguientes son equivalentes:

a) $AX = 0$ *tiene solución única* $X = 0$.

b) A es invertible.

c) det $A \neq 0$.

d) Las columnas de A forman una base para $\mathbb{R}^n$. *También las filas de A forman una base para* $\mathbb{R}^n$.

e) La dimensión del espacio de soluciones del sistema homogéneo $AX = 0$ *es igual a cero.*

Ejercicios 4

1) Calcule la dimensión del espacio-solución de los siguientes sistemas de ecuaciones lineales:

a) $$\begin{aligned} x - 3y + z &= 0 \\ 2x + y - 2z &= 0 \\ x + 4y - 4z &= 0 \end{aligned}$$

b) $$\begin{aligned} x + 2y &= 0 \\ 3x + 6y &= 0 \\ 5x + 9y &= 0 \end{aligned}$$

c) $$\begin{aligned} 4x + 2y - 3z &= 0 \\ 6x + 3y - 5z &= 0 \\ x + y + 2z &= 0 \end{aligned}$$

d) $$\begin{aligned} 2x + 2y - z &= 0 \\ x + y + z &= 0 \\ 2x - 4y + 3z &= 0 \end{aligned}$$

e) $$\begin{aligned} x - 3y + 4z &= 0 \\ 3x - 5y + 5z &= 0 \end{aligned}$$

f) $$\begin{aligned} -2x + y + 7z &= 0 \\ 4x + 9y + 2z &= 0 \end{aligned}$$

2) Para cada uno de los seis espacios-solución del ejercicio 1 anterior, calcule, si existe, una base ortonormal.

5. Contexto económico

a. El análisis insumo-producto de Leontief (1936)

El *análisis insumo-producto* (1936b)) del economista (entonces soviético) Wassily Leontief [1906–1999], y que fuera la segunda gran herramienta de la economía lineal en aparecer después del modelo Walras-Cassel, constituye otra adaptación de la teoría walrasiana del equilibrio general al estudio de la interdependencia cuantitativa que existe entre algunas actividades económicas. Está basado en la idea de que *una parte muy considerable del esfuerzo de una economía moderna está dedicada a la producción de bienes intermedios*, y que se encuentran muy ligados al producto final. Así, un cambio en el nivel de producción de un bien final (digamos, una casa) implica cambios en la producción de los bienes intermedios asociados a su producción (cemento, acero, vidrios, etc.) y, a su vez, en los insumos utilizados para la producción de estos insumos intermedios, etc.

Leontief inicialmente estudió una *economía cerrada* (es decir, donde todos los bienes eran intermedios, siendo los consumibles también bienes intermedios en la producción de servicios y otros bienes). Buscaba hallar un estado de *equilibrio* en el que sólo lo justo de cada bien se produjera para satisfacer los requerimientos de insumos de todos los otros bienes. De esta manera, también podría identificar los *precios de equilibrio* de los bienes. Posteriormente, el énfasis de Leontief se centró ya no en una economía cerrada sino en una economía en la que la demanda final estuviera exógenamente determinada. Entonces encontraba los niveles de actividad de los distintos sectores de la economía consistentes con esta demanda (incluyendo niveles de empleo). Era el modelo de *economía abierta*.

Leontief consideraba una economía en la cual bienes tales como hierro, carbón, algodón, etc., se producen en sus respectivas industrias mediante un insumo primario como la mano de obra, y por medio de insumos tales como hierro, carbón, algodón, etc. Observemos cómo Leontief rechaza la idea de que ciertas industrias son etapas anteriores de la producción y, otras, posteriores. Así, se opone a la idea de que inevitablemente se debe encontrar una industria (tal como la agricultura) que sólo le vende a otra (como la manufactura) pero que no compra nada de ésta. Niega que uno pueda seguir la fabricación de un pan desde las primeras etapas a través de una *jerarquía unidireccional* de industrias. Para Leontief, el mundo real es el de *relaciones interindustriales multidireccionales*.

Industria	Insumos agricultura	Insumos manufactura	Demanda final	Producciones finales
Agricultura	75	100	125	300
Manufactura	40	40	80	160
Mano de obra	25	45	0	70

Figura 7. Matriz insumo-producto simplificada

Como ejemplo de todo esto, supongamos una economía muy simplificada en la que sólo hay dos industrias: agricultura y manufactura. Cada una requiere, directamente, mano de obra y también utiliza elementos de la otra industria en su proceso productivo. La figura 7 muestra una forma simplificada de la economía.

Allí, en la primera fila, de las 300 unidades de producción agrícola, 125 unidades van al consumo final (hogares y gobierno), 100 unidades a insumos para la industria de manufactura y 75 unidades a insumos para la industria agrícola. La segunda fila es similar. En la tercera fila aparece que de 70 unidades (horas-hombre) de mano de obra, 45 serán requeridas por la industria de la manufactura y 25 por la industria agrícola. La cantidad de mano de obra está dada exógenamente y, en este ejemplo, es considerada únicamente como insumo; sin embargo, la mano de obra podría ser considerada como una industria más en el modelo. Como se puede ver, las matrices insumo-producto son sólo una forma fácil de organizar información sobre ciertas transacciones del sistema económico. Es una útil tabulación en donde grandes y complicadas economías pueden describirse mediante ciertos números que, en muchas ocasiones, son posibles de encontrar o de estimar.

¿Cuáles son las hipótesis implícitas en el análisis insumo-producto de Leontief? Para ver esto, transformemos la figura 7 en una más descriptiva y general (figura 8).

	Insumos industria 1	Insumos industria 2	Demanda final	Producción total
Industria 1	x_{11}	x_{12}	c_1	$\bar{x}_1$
Industria 2	x_{21}	x_{22}	c_2	$\bar{x}_2$
Mano de obra	x_{01}	x_{02}	0	$\bar{x}_0$

Figura 8. Matriz insumo-producto

donde x_{1i} indica la cantidad del insumo 1 utilizado por la industria i (para $i = 1, 2$); y de forma similar, x_{2i} y x_{0i} indican, respectivamente, la cantidad del insumo 2 y la cantidad de mano de obra utilizada por cada una de las industrias.

De esta tabla podemos escribir las *funciones de producción* como

$$\begin{aligned} \bar{x}_1 &= F_1\,(\,x_{11},\,x_{21},\,x_{01}\,) \\ \bar{x}_2 &= F_2\,(\,x_{12},\,x_{22},\,x_{02}\,) \end{aligned} \tag{1}$$

donde

$$\begin{aligned} x_{11}+x_{12}+c_1 &= \bar{x}_1 \\ x_{21}+x_{22}+c_2 &= \bar{x}_2 \\ x_{01}+x_{02} &= \bar{x}_0 \end{aligned} \tag{2}$$

En este punto, *Leontief asume, explícitamente, que las funciones de producción tienen rendimientos constantes a escala* y que toman una cierta cantidad *mínima* de cada insumo para producir una unidad de producto. Es decir, asume que

$$\begin{aligned} \bar{x}_1 &= \text{mín}\left(\frac{x_{11}}{a_{11}},\frac{x_{21}}{a_{21}},\frac{x_{01}}{a_{01}}\right) \\ \bar{x}_2 &= \text{mín}\left(\frac{x_{12}}{a_{12}},\frac{x_{22}}{a_{22}},\frac{x_{02}}{a_{02}}\right) \end{aligned} \tag{3}$$

donde a_{ij} es la cantidad mínima de insumo i que se requiere para producir una unidad de producto j (a éstas se les conocerá en la literatura como *funciones de producción Leontief*). Para nuestro caso particular de la figura 7, tendremos que (sabiendo que por las igualdades (3), podemos asumir $a_{ij}=\dfrac{x_{ij}}{\bar{x}_j}$) $a_{11}=\dfrac{x_{11}}{\overline{x}_1}=\dfrac{75}{300}=0.25$; $a_{12}=\dfrac{x_{12}}{\overline{x}_2}=\dfrac{100}{160}=0.625$; $a_{21}=0.13$; $a_{22}=0.25$; $a_{01}=0.08$; $a_{02}=0.28$. Así, las funciones de producción (según Leontief) para aquella economía particular de la figura 7, son

$$\begin{aligned} \bar{x}_1 &= \text{mín}\left(\frac{x_{11}}{0.25},\frac{x_{21}}{0.13},\frac{x_{01}}{0.08}\right) \\ \bar{x}_2 &= \text{mín}\left(\frac{x_{12}}{0.625},\frac{x_{22}}{0.25},\frac{x_{02}}{0.28}\right) \end{aligned}$$

En general, de (2), se tiene que

$$[\overline{x}_1\ ,\ \overline{x}_2]^T=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} \overline{x}_1 \\ \overline{x}_2 \end{bmatrix}+[c_1\ ,\ c_2]^T$$

ó

$$\overline{X}=A\overline{X}+C$$

donde $\overline{X} = [\overline{x}_1\ ,\ \overline{x}_2]^T$, $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $C = [c_1\ ,\ c_2]^T$ y, por tanto,

$$\overline{X} = (I_2 - A)^{-1}C. \tag{4}$$

Sin embargo, no es claro que una relación como la anterior también se dé para la mano de obra. Recordemos que de (2), $x_{01} + x_{02} = \overline{x}_0$ y que, por tanto, como $x_{01} = a_{01}\overline{x}_1$ y $x_{02} = a_{02}\overline{x}_2$, entonces $x_0 = a_{01}\overline{x}_1 + a_{02}\overline{x}_2$, y así, $\overline{x}_0 = [a_{01}\ ,\ a_{02}] \cdot (I_2 - A)^{-1}C$ donde $(\cdot)$ es el producto interno entre los vectores $[a_{01}\ ,\ a_{02}]$ y $(I_2 - A)^{-1}C$.

¿Y qué acerca de los precios en el modelo de Leontief? *Él asume que, en equilibrio, los precios deben igualar los costos por unidad.* Así, si w es el salario por hora-hombre, entonces

$$\begin{aligned} p_1 &= a_{11}p_1 + a_{21}p_2 + a_{01}w \\ p_2 &= a_{12}p_1 + a_{22}p_2 + a_{02}w \end{aligned}$$

o

$$p = A^T p + wA_0$$

donde $p = [\,p_1, p_2\,]^T$, $A^T = [\begin{smallmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{smallmatrix}]$, $A_0 = [\,a_{01}, a_{02}\,]^T$ y resolviendo este sistema matricial en términos del numerario "salario" (w) obtenemos[3]

$$p = w(\,I_2 - A^T\,)^{-1}A_0 \tag{5}$$

Por lo tanto, mediante cualquier método de los estudiados en la lección 3 para el cálculo de matrices inversas, obtenemos que

$$p_1 = \frac{a_{01}(\,1 - a_{22}\,) + a_{02}a_{12}}{(\,1 - a_{11}\,)(\,1 - a_{22}\,) - a_{12}a_{21}}\,w$$

$$p_2 = \frac{a_{02}(\,1 - a_{11}\,) + a_{01}a_{21}}{(\,1 - a_{11}\,)(\,1 - a_{22}\,) - a_{12}a_{21}}\,w$$

Ahora: observemos que para que p_1, p_2 sean precios mayores que cero (y para que las soluciones de (4) sean positivas) debemos tener la llamada condición Hawkins-Simon[4]:

$$(\,1 - a_{11}\,)(\,1 - a_{22}\,) > a_{12}a_{21} \tag{6}$$

que es la más sutil restricción del modelo de Leontief básico. ¿Qué significa? Veamos esto.

---

[3] Notemos que aquí también sólo podemos encontrar *precios relativos* debido al tipo de funciones de producción escogido por Leontief.

[4] Hawkins D. y H. A. Simon (1949), Some Conditions of Macroeconomics Stability, *Econometrica*, vol. 17, 245-248.

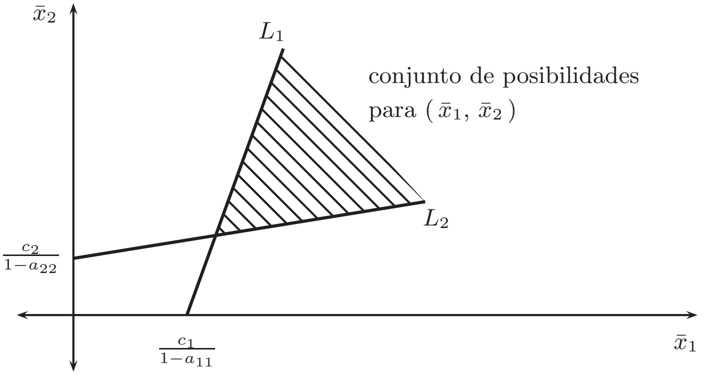


Figura 9

En la figura 9 dibujamos las restricciones $\bar{x}_1 = a_{11}\bar{x}_1 + a_{12}\bar{x}_2 + c_1$ (recta L_1), y $\bar{x}_2 = a_{21}\bar{x}_1 + a_{22}\bar{x}_2 + c_2$ (recta L_2). Si una economía busca producir un par de determinadas demandas finales c_1 y c_2, ¿qué condiciones sobre los a_{ij} deben tenerse para que existan las *cantidades positivas* $\bar{x}_1$, $\bar{x}_2$ correspondientes? De la figura 9 se ve claramente que el requisito es que la pendiente de la recta L_2 sea menor que la pendiente de la recta L_1; es decir,

$$\frac{1-a_{11}}{a_{12}} > \frac{a_{21}}{1-a_{22}} \qquad \text{ó} \qquad (1-a_{11})(1-a_{22}) > a_{12}a_{21}$$

que es la condición (6). Así, *la condición (6) es una condición mínima para que el sistema de Leontief pueda funcionar como un sistema económico.*

Para nuestro caso particular que venimos estudiando, tendremos entonces que los precios, relativos al salario de la mano de obra, son

$$\frac{p_1}{w} = \frac{(0.08)(1-0.25) + (0.28)(0.625)}{(1-0.25)(1-0.25) - (0.625)(0.13)} = 0.49$$

$$\frac{p_2}{w} = \frac{(0.3)(1-0.1) + (0.04)(0.16)}{(1-0.1)(1-0.17) - (0.83)(0.16)} = 0.46$$

es decir, p_1, p_2 son, aproximadamente, la mitad del salario.

Utilizando el álgebra matricial es posible generalizar el modelo de Leontief a n industrias. Sean

$\bar{X} = [\,\bar{x}_1, \ldots, \bar{x}_n\,]^T$ (vector de cantidades totales de producción)

$X_0 = [\,x_{01}, \ldots, x_{0n}\,]$ (vector de horas de mano de obra para cada industria)

$A = [\,a_{ij}\,]_{i,j=1,\ldots,n}$ (matriz de proporciones fijas de producción)

$A_0 = [\,a_{01}, \ldots, a_{0n}\,]^T$ (vector de proporciones fijas de mano de obra)

$C = [\,c_1, \ldots, c_n\,]^T$ (vector de consumos finales)

$P = [\,p_1, \ldots, p_n\,]^T$ (vector de precios)

Entonces el sistema estará definido por las ecuaciones matriciales

$$\bar{X} = (\,I_n - A\,)^{-1}C \tag{4'}$$

(si la inversa de $(\,I_n - A\,)$ existe, donde I_n es la matriz identidad $n \times n$)

$$P = w\,(\,I_n - A^T\,)^{-1}A_0 \tag{5'}$$

$$x_{01} = a_{01}\overline{x}_1,\quad x_{02} = a_{02}\overline{x}_2,\ \ldots,\ x_{0n} = a_{0n}\overline{x}_n$$

Los sistemas explícitos correspondientes a las ecuaciones matriciales en $(4')$ y $(5')$ son:

$$\begin{aligned}\begin{bmatrix}\overline{x}_1\\ \overline{x}_2\\ \vdots\\ \overline{x}_n\end{bmatrix} &= \left(\begin{bmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \cdots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix} - \begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \cdots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{bmatrix}\right)^{-1}\begin{bmatrix}c_1\\ c_2\\ \vdots\\ c_n\end{bmatrix} && (4')\\ &= \begin{bmatrix}1-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & 1-a_{22} & \cdots & -a_{2n}\\ \vdots & \vdots & \cdots & \vdots\\ -a_{n1} & -a_{n2} & \cdots & 1-a_{nn}\end{bmatrix}^{-1}\begin{bmatrix}c_1\\ c_2\\ \vdots\\ c_n\end{bmatrix} && (L1)\end{aligned}$$

$$\begin{aligned}\begin{bmatrix}p_1\\ p_2\\ \vdots\\ p_n\end{bmatrix} &= w\left(\begin{bmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \cdots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix} - \begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \cdots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{bmatrix}\right)^{-1}\begin{bmatrix}a_{01}\\ a_{02}\\ \vdots\\ a_{0n}\end{bmatrix} && (5')\\ &= w\begin{bmatrix}1-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & 1-a_{22} & \cdots & -a_{2n}\\ \vdots & \vdots & \cdots & \vdots\\ -a_{n1} & -a_{n2} & \cdots & 1-a_{nn}\end{bmatrix}^{-1}\begin{bmatrix}a_{01}\\ a_{02}\\ \vdots\\ a_{0n}\end{bmatrix} && (L2)\end{aligned}$$

Nuevamente, aquí, nos enfrentamos al problema de garantizar que en las dos ecuaciones matriciales anteriores, siempre exista una solución no-negativa. Y la respuesta, que presentamos en el teorema general siguiente, es afirmativa bajo ciertos requisitos: son las *condiciones Hawkins-Simon* que habíamos estudiado anteriormente, en su versión más elemental.

Teorema 10. ***(Condiciones Hawkins- Simon)***
Si las entradas de la matriz

$$\begin{bmatrix} 1-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & 1-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & 1-a_{nn} \end{bmatrix}$$

satisfacen

$$1-a_{11} > 0, \quad \begin{vmatrix} 1-a_{11} & -a_{12} \\ -a_{21} & 1-a_{22} \end{vmatrix} > 0$$

$$\begin{vmatrix} 1-a_{11} & -a_{12} & -a_{13} \\ -a_{21} & 1-a_{22} & -a_{23} \\ -a_{31} & -a_{32} & 1-a_{33} \end{vmatrix} > 0, \ \cdots, \ \begin{vmatrix} 1-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & 1-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & 1-a_{nn} \end{vmatrix} > 0$$

entonces ambos sistemas, (L1) y (L2), del modelo Leontief tienen una única solución no-negativa.

Demostración[5]
Probaremos esto para el sistema (L1). El caso para el sistema (L2) es similar e inmediato a partir del caso (L1).
Consideremos entonces el sistema general (L1) original

$$\begin{array}{lllll} (1-a_{11})\overline{x}_1 & -a_{12}\overline{x}_2 - a_{13}\overline{x}_3 & -\cdots & -a_{1n}\overline{x}_n = c_1 & \\ -a_{21}\overline{x}_1 & +(1-a_{22})\overline{x}_2 - a_{23}\overline{x}_3 & -\cdots & -a_{2n}\overline{x}_n = c_2 & \\ -a_{31}\overline{x}_1 & -a_{32}\overline{x}_2 + (1-a_{33})\overline{x}_3 & -\cdots & -a_{3n}\overline{x}_n = c_3 & (*) \\ \vdots & \vdots & \vdots & \cdots \quad \vdots & \vdots \\ -a_{n1}\overline{x}_1 & -a_{n2}\overline{x}_2 - a_{n3}\overline{x}_3 & -\cdots & +(1-a_{nn})\overline{x}_n = c_n & \end{array}$$

Y vamos a probar el resultado por el método de inducción matemática sobre los números naturales $\mathbb{N}$ (Volumen 0 (Fundamentos)). El caso $n = 1$, es $(1 - a_{11})\overline{x}_1 = c_1$ y esto implica que $\overline{x}_1 = \dfrac{c_1}{1-a_{11}} \geq 0$ pues $c_1 \geq 0$, $1 - a_{11} > 0$, por hipótesis. Por lo tanto, el teorema es cierto para $n = 1$.

Ahora asumamos el teorema cierto para $n-1$ y probémoslo también para n. Para ello, notamos que el sistema $(*)$ anterior se puede reducir, mediante

[5] Esta prueba sigue la muy elegante demostración presentada en Nikaido, H. (1968), *Convex Structures and Economic Theory*, New York: Academic Press.

eliminación gaussiana, a un sistema de la forma

$$\begin{array}{llll} (1-a_{11})\overline{x}_1 & -a_{12}\overline{x}_2\cdots & -a_{1n}\overline{x}_n = c_1 & \\ & b_{22}\overline{x}_2\cdots & -b_{2n}\overline{x}_n = c_2 & (**) \\ & \cdots \quad \vdots & \vdots & \\ & b_{n2}\overline{x}_2\cdots & -b_{nn}\overline{x}_n = c_n & \end{array}$$

donde los $b'_{ij}s$ resultan de sumar, a la i-ésima ecuación ($2 \leq i \leq n$), la primera ecuación multiplicada por $-a_{i1}/(1-a_{11})$. Con un poco de cuidado se puede observar que la relación entre los cofactores del sistema original $(*)$ con los nuevos de $(**)$ es, para $2 \leq k \leq n$,

$$\begin{vmatrix} b_{22} & \cdots & b_{2k} \\ \vdots & \cdots & \vdots \\ b_{k2} & \cdots & b_{kk} \end{vmatrix} = \frac{1}{1-a_{11}} \begin{vmatrix} 1-a_{11} & -a_{12} & \cdots & -a_{1k} \\ -a_{21} & 1-a_{22} & \cdots & -a_{2k} \\ \vdots & \cdots & \cdots & \vdots \\ -a_{k1} & -a_{k2} & \cdots & 1-a_{kk} \end{vmatrix}$$

Dada la hipótesis inicial, los determinantes de la derecha son todos positivos y, por tanto, los de la izquierda también lo son. Así el sistema de ecuaciones nuevo (desde la ecuación 2 hasta la ecuación n) satisface las condiciones de Hawkins-Simon, y la hipótesis de inducción nos asegura que este sistema tiene una solución no-negativa $\overline{x}_2, \overline{x}_3, \cdots, \overline{x}_n$. Pero como de la primera ecuación de $(**)$ se tiene

$$\overline{x}_1 = [c_1 + a_{12}\overline{x}_2 + a_{13}\overline{x}_3 + \cdots + a_{1n}\overline{x}_n]/(1-a_{11})$$

entonces también se tendrá que $\overline{x}_1$ es no-negativa. Para terminar, sólo basta recordar que la solución del sistema $(**)$ es la misma solución del sistema $(*)$. ■

Así finalizamos esta breve exposición del modelo "estático" de Leontief, aunque cabe advertir que existen modelos más avanzados de él que incluyen características temporales como flujos de bienes e inventarios de capital. Aún así, en muchos casos, los intentos de generalización del modelo básico de Leontief han conducido a que se asimilen casi completamente a los sistemas walrasianos. Obviamente, debe hacerse aquí la invitación al lector al estudio de la obra original de Leontief. Existe una versión en castellano que reúne algunos de sus principales trabajos sobre la matriz insumo-producto.[6]

[6] Leontief, W. (1993), *Análisis económico input-output*, Madrid: Editorial Planeta.

Nota 10. (Sobre Wassily Leontief)
Luego de llegar a la conclusión de que su llamado "análisis de equilibrio parcial" no podía proveer las bases suficientes para entender la estructura y funcionamiento del sistema económico, en 1931 Leontief formuló su propia teoría del equilibrio general con posibilidades de implementación empírica. Fue así como, en 1932, recibió una concesión de investigación por la compilación de la primera tabla *input-output* de la economía estadounidense. Esta tabla estuvo inspirada en el análisis de sistemas de producción lineales, que sirvieron como instrumento en el desarrollo de la teoría neowalrasiana moderna. También fue crucial para revivir la clásica teoría ricardiana y fue empleada por Piero Sraffa y los neorricardianos en los años sesenta para revivir las teorías de Ricardo y Marx.

Pero las contribuciones de Leontief no se limitaron a su tabla *input-output.* Su artículo de 1936[7] sobre *mercancías compuestas* atacaba el problema microeconómico de agregación (utilizado, en ocasiones, para justificar el análisis de equilibrio parcial). De igual forma, sus primeras revisiones sobre la *Teoría General* de Keynes se convirtieron en pasos importantes para la síntesis neokeynesiana sobre salarios nominales fijos en la interpretación de la teoría de Keynes. Su artículo de 1933[8] sobre el análisis del comercio internacional es aún importante, y su contribución de 1946[9] sobre los contratos de trabajo es ahora una aplicación clásica del modelo principal-agente de la teoría de juegos con información asimétrica. Una de las más importantes contribuciones fue su estudio de 1953 donde encontró que las exportaciones americanas eran bienes intensivos en trabajo y no en capital (conocida como la "paradoja de Leontief") y que sirvió, según algunos, para validar la teoría convencional del comercio internacional.

Leontief fue un matemático empírico apasionado por los datos. Sus últimos escritos contienen numerosos comentarios sobre las dificultades de teorizar a priori en economía, sobre la falta de atención de los economistas a la calidad de las estadísticas que usan en los trabajos empíricos, sobre la escasez de investigación en las técnicas econométricas y, sobre todo, en la necesidad de invertir en una adecuada recolección de datos si se desea alcanzar un mejor entendimiento de la vida real.

[7] Leontief, W. (1936), Composite Commodities and the Problem of the Index Numbers, *Econometrica*, vol. 4 (1), 39-59.
[8] Leontief, W. (1933), The Use of Indifference Curves in the Analysis of Foreign Trade, *Quarterly Journal of Economics*, vol. 97, 493-503.
[9] Leontief, W. (1946), The Pure Theory of the Guaranteed Annual Wage Contract, *Journal of Political Economy*, vol. 54, 99-150.

Ejercicios complementarios

1) ¿Cuáles de los siguientes conjuntos son espacios vectoriales?:

   a) $\mathbb{R}^\infty = \{(x_1, x_2, \ldots, x_n, \ldots) / x_n \in \mathbb{R}\}$ con suma y producto por escalar a la manera usual.
   b) $W_1 \subseteq \mathbb{R}^\infty$ definido por $W_1 = \{(k, 0, k, 0, k, \ldots) / k \in \mathbb{R}\}$
   c) $W_2 \subseteq \mathbb{R}^\infty$ definido por $W_2 = \{(x_1, x_2, \ldots, x_n, \ldots) / x_{t+1} \leq x_t$ para $t = 1, 2, \ldots\}$
   d) $W_3 \subseteq \mathbb{R}^3$ definido por $W_3 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 / x_1 + x_2 + x_3 = 0\}$
   e) $W_4 \subseteq \mathbb{R}^3$ definido por $W_4 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 / x_1 + x_2 + x_3 = 1\}$

2) ¿El conjunto $\{(a_1, \ldots, a_n) \in \mathbb{R}^n / a_1 a_2 \cdots a_n = 1\}$ es un subespacio de $\mathbb{R}^n$?

3) En el espacio vectorial $\mathcal{F}_\mathbb{R}$ de todas las funciones $f : \mathbb{R} \longrightarrow \mathbb{R}$, ¿cuáles de los siguientes conjuntos son subespacios?:

   a) $\{f : \mathbb{R} \longrightarrow \mathbb{R} / f(x) = mx + n$ para $m, n \in \mathbb{Z}$ fijos$\}$
   b) $\{f : \mathbb{R} \longrightarrow \mathbb{R} / f(x) = ax^2 + bx + c$ para $a, b, c \in \mathbb{R}$ fijos$\}$

4) En el espacio de las matrices $n \times n$, $\mathfrak{M}_{n\times n}$, ¿cuáles de los siguientes conjuntos son subespacios?:

   a) $\{A \in \mathfrak{M}_{n\times n} / AA^T = I_{n\times n}\}$
   b) $\{A \in \mathfrak{M}_{n\times n} / A$ es invertible$\}$
   c) $\{A \in \mathfrak{M}_{n\times n} / A^T = A\}$

5) Exprese explícitamente los siguientes espacios vectoriales:

   a) $\left\langle \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix} \right\rangle$ en $\mathfrak{M}_{2\times 2}$
   b) $\langle (1, 1), (0, 0) \rangle$ en $\mathbb{R}^2$
   c) $\langle (1, 1) \rangle + \langle (2, 1) \rangle$ en $\mathbb{R}^2$
   d) $\langle (1, 1, 1) \rangle + \langle (2, 1, 3) \rangle$ en $\mathbb{R}^3$
   e) $\langle e^x, 1 + x \rangle$ en $\mathcal{F}_\mathbb{R}$

6) ¿Será que en el espacio $\mathcal{F}_\mathbb{R}$ de todas las funciones $f : \mathbb{R} \longrightarrow \mathbb{R}$, la función $f(x) = e^x + x + 1$ es combinación lineal de

$$f_1(x) = e^x, \quad f_2(x) = x - e^x, \quad f_3(x) = 3x + 2\,?$$

7) a) En el espacio de las matrices $\mathfrak{M}_{3\times 3}$, ¿será que la matriz

$$A = \begin{bmatrix} 2 & 11 & -5 \\ -2 & -3 & \frac{3}{2} \\ \frac{17}{3} & -4 & -14 \end{bmatrix}$$

es combinación lineal de las matrices

$$A_1 = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 2 & 0 \\ 4 & 1 & -3 \end{bmatrix}, \; A_2 = \begin{bmatrix} -4 & 5 & 1 \\ -2 & 3 & \frac{1}{2} \\ \frac{1}{3} & 0 & 8 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 2 & -6 & 5 \\ 7 & 2 & -1 \\ 1 & 3 & 0 \end{bmatrix}?$$

b) En el espacio de las matrices $\mathfrak{M}_{4\times 4}$, ¿será que la matriz

$$A = \begin{bmatrix} 38 & -50 & -61 & -109 \\ -25 & 39 & -73 & -14 \\ 22 & 7 & 20 & 27 \\ 52 & 22 & 74 & 105 \end{bmatrix}$$

es combinación lineal de las matrices A_1, A_2 y A_3, donde

$$A_1 = \begin{bmatrix} 12 & -6 & -7 & -1 \\ 4 & 5 & 5 & 3 \\ 2 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -10 & -8 & 4 & 1 \\ -11 & 6 & 0 & 0 \\ 0 & 1 & 2 & 1 \\ 3 & 4 & 5 & 2 \end{bmatrix},$$

$$A_3 = \begin{bmatrix} 3 & 0 & 8 & 7 \\ 0 & 4 & 0 & 9 \\ 8 & 0 & 10 & 9 \\ 3 & 1 & 2 & 0 \end{bmatrix}?$$

8) Encuentre la dimensión de:

a) El hiperplano de $\mathbb{R}^4$, $\{\,(x, y, z, w)\,/\,x + y + z + w = 0\,\}$

b) El espacio-solución de $AX = 0$, donde

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{y} \quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

9) Encuentre una base y la dimensión del espacio de todas las matrices simétricas 2×2.

10) Encuentre una base y la dimensión para la intersección de los planos $x + 2y + 3z = 0$ y $2x - y - z = 0$. Geométricamente, ¿qué significa esto?

11) Elabore un cuadro que muestre los distintos espacios vectoriales estudiados en la presente lección (incluidos los de los ejercicios) y sus correspondientes dimensiones. Si son infinito-dimensionales, señalarlo así.

12) Describa geométricamente el subespacio de $\mathbb{R}^3$ generado por

    a) $(0,1,0)$, $(0,0,0)$, $(0,5,0)$
    b) $(0,1,1)$, $(0,2,4)$, $(0,0,3)$
    c) Los seis vectores anteriores

13) ¿Será que $(1,2,3)$ es generado por $(1,1,2)$, $(1,0,0)$ y $(2,2,4)$? ¿Y por $(1,1,2)$, $(1,0,0)$ y $(2,2,0)$?

14) Utilizando el método gaussiano, pruebe que una base para el espacio generado por los vectores $(1,2,-3,4)$, $(-2,1,-7,-18)$, $(2,5,-3,4)$, $(2,10,-2,7)$ es

$$\beta = \left\{ \left(1,0,0,-\frac{39}{14}\right), \left(0,1,0,\frac{13}{14}\right), \left(0,0,1,-\frac{23}{14}\right) \right\}$$

15) Encuentre dos bases diferentes para el subespacio de $\mathbb{R}^3$ definido por

$$\{ (x,y,z) \in \mathbb{R}^3 \,/\, y = z \}$$

16) ¿Será cierto que si $\{\beta_1, \beta_2, \beta_3\}$ forma una base para $\mathbb{R}^3$ y W es un subespacio de $\mathbb{R}^3$, entonces algún *subconjunto* de $\{\beta_1, \beta_2, \beta_3\}$ forma una base para W?

17) ¿Serán las columnas de la matriz

$$\begin{bmatrix} 1 & 5 & 2 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{bmatrix}$$

linealmente independientes? La misma pregunta para las filas de la matriz.

18) Pruebe que si W_1, $W_2 \subseteq \mathbb{R}^5$ son subespacios de dimensión 3, entonces $W_1 \cap W_2$ es un subespacio *no-nulo* de $\mathbb{R}^5$.

19) Pruebe que si A es una matriz $n \times n$ tal que $A^2 = A$ y sus columnas son linealmente independientes, entonces $A = I_n$ (la matriz identidad). [Indicación: pruebe que A es invertible].

20) Pruebe que si x, y, $z \in \mathbb{R}^n$ son linealmente independientes, entonces también x, $x+y$, $x+y+z$ son linealmente independientes.

21) Pruebe que $|x \cdot y| = ||\,x\,||\,||\,y\,||$ si, y sólo si, x y y son linealmente dependientes.

22) Pruebe que si $||\,x + y\,|| = ||\,x\,|| + ||\,y\,||$, entonces x y y son linealmente dependientes. ¿Será cierto el recíproco? ¿Por qué?

23) En $\mathcal{F}_{\mathbb{R}}$ muestre que los siguientes pares de funciones son linealmente independientes:

a) $1, x$	b) $\operatorname{sen} x, \operatorname{sen} 2x$	c) $x, \operatorname{sen} x$
d) x, x^2	e) x, x^3	f) $\cos x, \cos 3x$

24) ¿Será cierto que si $\beta = \{\beta_1, \beta_2, ..., \beta_n\}$ es una base para $\mathbb{R}^n$ entonces cualquier conjunto β' formado por n vectores que son combinaciones lineales (no nulas) de los vectores de β también conforma una base para $\mathbb{R}^n$?

25) Encuentre una base para las soluciones al sistema

$$\begin{bmatrix} 1 & 2 & 1 \\ 3 & 4 & 2 \\ 2 & 4 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

¿Por qué no preguntamos por una base para las soluciones al sistema

$$\begin{bmatrix} 1 & 2 & 1 \\ 3 & 4 & 2 \\ 2 & 4 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}?$$

26) Encuentre una base para la intersección de los planos

$$x + 2y + z = 1 \quad \text{y} \quad 3x + y - 2z = 0$$

27) Considere una economía de tipo Leontief de dos sectores cuya estructura de producción se puede resumir por medio de una matriz insumo-producto. La matriz de coeficientes técnicos es

$$A = \begin{bmatrix} 0.2 & 0.4 \\ 0.3 & 0.1 \end{bmatrix}$$

Si la demanda final es

$$C = \begin{bmatrix} 30 \\ 40 \end{bmatrix}$$

encuentre la producción necesaria para satisfacer esta demanda (Indicación: calcule $\overline{X} = (I_2 - A)^{-1}C$).

28) En una economía de tipo Leontief con matriz de producción

$$A = \begin{bmatrix} 0.20 & 1.28 & 0.25 \\ 0.01 & 0.14 & 0.02 \\ 0.01 & 0.14 & 0.12 \end{bmatrix}$$

y vector de consumos (o demandas) finales

$$C = \begin{bmatrix} 10 \\ 20 \\ 12 \end{bmatrix}$$

a) ¿Cuál será el vector de cantidades totales de producción?
b) ¿Y cuál el vector de precios unitarios?
c) Si el vector de cantidades iniciales de mano de obra es $A_0 = [1, 2, 3]^T$, ¿cuál es el vector de cantidades de mano de obra asignadas a cada industria?

29) Construya las curvas de nivel de la función de producción Leontief $\overline{x} = \text{mín}\{\frac{x_{11}}{0.1}, \frac{x_{21}}{0.16}\}$. De manera similar para $\overline{x} = \text{mín}\{\frac{x_{11}}{a}, \frac{x_{21}}{b}\}$ con $a, b > 0$ conocidos.

30) Suponga una economía de tipo Leontief en la que hay dos industrias y cada una de ellas utiliza, como insumos, elementos de su propia industria y de la otra. Si los coeficientes técnicos son $a_{11} = 0.25$, $a_{12} = 0.02$, $a_{21} = 0.04$ y $a_{22} = 0.15$, y las demandas finales son $c_1 = 150$ y $c_2 = 200$, ¿cuál es la producción de cada industria?

31) Suponga una economía de tipo Leontief descrita por la siguiente matriz insumo-producto:

	Agricultura	Manufactura	Servicios	Demanda final
Agricultura	35	5	6	39
Manufactura	5	62	23	60
Servicios	10	26	42	131

a) Encuentre la matriz A de coeficientes técnicos.
b) Encuentre los niveles de producción de equilibrio.
c) Suponga que las demandas finales cambian de forma tal que

$$C = \begin{bmatrix} 50 \\ 60 \\ 132 \end{bmatrix}$$

¿Cuáles son los nuevos niveles de producción de equilibrio? ¿Cómo debe cada sector proveer la producción extra?

32) Una economía produce tres bienes, los que para ser producidos requieren de dos tipos de trabajo: uno *calificado* y otro *no-calificado*. Los requerimientos por unidad de producto se presentan en la siguiente matriz:

$$B = \begin{bmatrix} 5 & 5 & 4 \\ 4 & 8 & 10 \end{bmatrix}$$

La primera fila representa los requerimientos de trabajo calificado y la segunda los de trabajo no-calificado. Además, la producción requiere de insumos, los cuales se pueden representar por la siguiente matriz de coeficientes técnicos:

$$A = \begin{bmatrix} 0.08 & 0.12 & 0.10 \\ 0.15 & 0.20 & 0.15 \\ 0.10 & 0.20 & 0.10 \end{bmatrix}$$

La demanda de estos bienes es realizada por dos grupos de consumidores: los trabajadores no-calificados y los trabajadores calificados. Las funciones de demanda de los trabajadores no-calificados son

$$x_{11} = \frac{4w_1}{p_1}, \quad x_{12} = \frac{8w_1}{p_2}$$

donde x_{11} es la demanda que hacen estos consumidores del bien 1, p_1 es el precio de dicho bien, w_1 es el salario de los trabajadores no-calificados y p_2 el precio del bien 2. Este tipo de consumidores no demanda el bien 3.

Las funciones de demanda de los trabajadores calificados son

$$x_{22} = \frac{12w_2}{5p_2}, \quad x_{23} = \frac{24w_2}{5p_2}$$

donde x_{22} es la demanda que hace este tipo de consumidores del bien 2, x_{23} es la demanda del bien 3, w_3 es el salario de los trabajadores calificados y p_3 el precio del bien 3. Este tipo de consumidores no demanda el bien 1. Por último, la oferta de trabajo de los trabajadores no-calificados es 12 y la de los trabajadores calificados es 7.2.

a) Encuentre las funciones de demanda total de cada uno de los tres bienes.

b) Si el salario de los trabajadores no-calificados es 100 y el de los trabajadores calificados es 200, encuentre los precios tales que los beneficios de las empresas sean iguales a cero.

c) Dados estos precios, encuentre la demanda final, la producción necesaria para satisfacerla, y las demandas de trabajo calificado y no-calificado.

$*$33) ¿En dónde encuentra usted las diferencias y similitudes esenciales entre el modelo walrasiano de Cassel y el modelo de Leontief?

Lección 6

Transformaciones lineales

Introducción

Posterior al advenimiento de la teoría de vectores por parte de Lagrange, el siguiente paso en el desarrollo de la geometría analítica fue el impulso de la *teoría de las transformaciones afines* (del latín *affinis* que significa "relacionadas"). Contracciones de un plano en una línea, la elipse como resultado de la contracción del círculo y, en general, el estudio de las proyecciones de una figura geométrica y su relación con la figura original fueron objeto de estudio desde los problemas concretos de perspectiva de pintores y dibujantes, pasando por la fundación de la *geometría proyectiva* por Girard Desargues [1591–1661], su desarrollo posterior por Jean V. Poncelet [1788–1867] y Gaspar Monge [1746–1818], hasta su síntesis estructuralista en el siglo XX.

Una *transformación afín* del plano es aquélla bajo la cual un sistema coordenado dado Oxy, se transforma en otro sistema coordenado $O'x'y'$ donde los ejes x' y y' pueden tener una medida unitaria diferente, y el ángulo que forman entre ellos puede ser distinto al que forman x y y; además, debemos tener que el origen (O) se transforma en el otro origen (O') (figura 1).

Las transformaciones afines se utilizaban, fundamentalmente, para tres objetivos: primero, para resolver problemas geométricos tales como encontrar las propiedades que se preservan bajo transformaciones afines. Aquí, realizaban una transformación afín de la figura que querían estudiar en una mucho más simple (por ejemplo, para encontrar el triángulo con el área más pequeña que podía circunscribirse a una elipse, aplicaban una transformación afín a esta última reduciéndola a un círculo), allí estudiaban el problema correspondiente y luego, con la propiedad deseada, retornaban a la figura original.

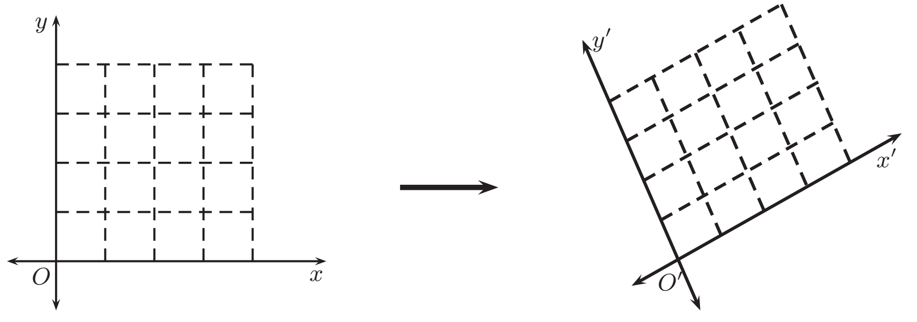


Figura 1

En segundo lugar, también las utilizaban para la clasificación de curvas y superficies de segundo orden. Aquí, el punto es que (como se puede demostrar) las elipses son afines unas a otras (es decir, podemos convertir una cualquiera en otra, mediante una transformación afín conveniente); también lo son las parábolas entre sí; y las hipérbolas entre sí. Sin embargo, no podemos convertir una elipse en una parábola, o una hipérbola en una parábola, mediante una transformación afín. Pareció entonces natural dividir todas las curvas de segundo orden en "clases afines" de curvas; es decir, dos curvas pertenecen a la misma clase afín si cualquiera de ellas puede transformarse, de forma afín, en la otra. Cuando se hizo esto, resultó que la reducción de una ecuación de segundo orden cualquiera, $Ax^2 + By^2 + Cxy + Dx + Ey + F = 0$, aparecía en una, y sólo una, de *nueve* clases afines. Estas son, para a, $b \neq 0$,

a) La elipse $\left(\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1\right)$

b) La elipse imaginaria $\left(\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = -1\right)$

c) El punto $\left(\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 0\right)$ (par de líneas paralelas que se intersectan en un punto lineal)

d) La hipérbola $\left(\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1\right)$

e) Un par de líneas que se intersectan $\left(\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 0\right)$

f) La parábola $(\, y^2 = 2ax \,)$

g) $x^2 - a^2 = 0$ (par de líneas paralelas)

h) $x^2 + a^2 = 0$ (par de líneas paralelas imaginarias)

i) $x^2 = 0$ (par de líneas rectas coincidentes)

Una tercera aplicación de las transformaciones afines fue a la teoría de transformaciones en "medios continuos" como, por ejemplo, en la teoría de la elasticidad, en la teoría de los campos magnéticos, en la de los campos eléctricos, etc. La idea fundamental aquí es que pequeños elementos de un medio continuo se transforman "casi" en forma afín. De allí la típica frase de que "en pequeñas dimensiones toda transformación es afín" (figura 2).

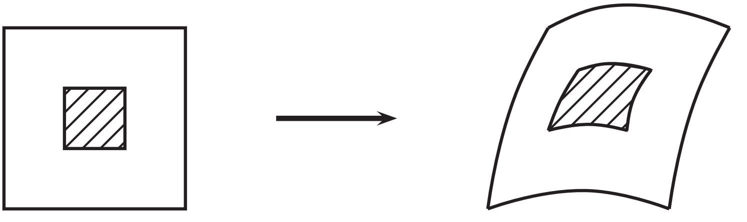


Figura 2

Ahora: para hallar fórmulas cartesianas de transformaciones afines explícitas, deduzcamos, por ejemplo, fórmulas mediante las cuales las coordenadas rectangulares cambian en forma afín. Sean x, y las coordenadas de un punto M relativo a los ejes Oxy. Traslademos estos ejes (paralelos a ellos mismos) a la posición $O'x'y'$, y sean ζ, η, las coordenadas del nuevo origen O' relativo a los "ejes viejos" x, y. Es claro, de la figura 3, que las nuevas coordenadas x', y' del punto M están conectadas con sus viejas coordenadas x, y, mediante las fórmulas de traslación paralela de ejes:

$$x = x' + \zeta$$
$$y = y' + \eta$$

o, en forma vectorial,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix} + \begin{bmatrix} \zeta \\ \eta \end{bmatrix}$$

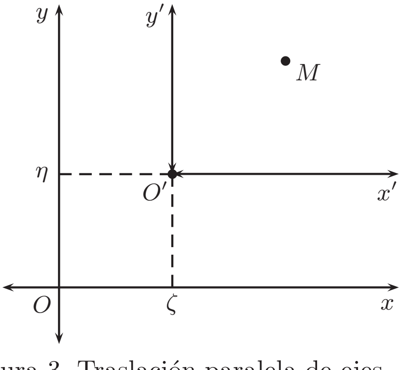


Figura 3. Traslación paralela de ejes

Ahora: si la transformación afín es rotar los ejes originales Oxy alrededor del origen (en forma contraria a las manecillas del reloj) un ángulo ϕ, entonces (con un poco de cuidado) podemos ver, de la figura 4, que

$$\begin{aligned} x &= x' \cos \phi - y' \operatorname{sen} \phi \\ y &= x' \operatorname{sen} \phi + y' \cos \phi \end{aligned}$$

o, en forma matricial,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \phi & -\operatorname{sen} \phi \\ \operatorname{sen} \phi & \cos \phi \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

Por tanto, una transformación afín que consista en una traslación de ejes y una rotación al sistema coordenado aparecería así:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \phi & -\operatorname{sen} \phi \\ \operatorname{sen} \phi & \cos \phi \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + \begin{bmatrix} \zeta \\ \eta \end{bmatrix}$$

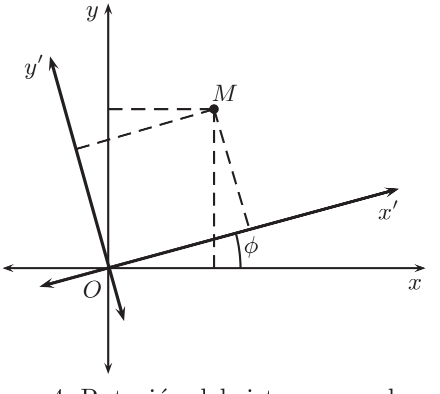


Figura 4. Rotación del sistema coordenado

A estos movimientos rígidos del plano se les llama *transformaciones ortogonales (afines) del plano.*

Si, además de esto, realizamos algún tipo de "contracción o dilatación de ejes", por ejemplo $kx' = x$, $ly' = y$, $(k, l > 0)$, entonces la transformación afín sería:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} kx' \\ ly' \end{bmatrix} = k \begin{bmatrix} x' \\ 0 \end{bmatrix} + l \begin{bmatrix} 0 \\ y' \end{bmatrix}$$

y una combinación de estos tres procedimientos básicos nos produce la *típica transformación afín del plano*:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} k \cos \phi & -l \operatorname{sen} \phi \\ k \operatorname{sen} \phi & l \cos \phi \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + \begin{bmatrix} \zeta \\ \eta \end{bmatrix} \tag{1}$$

es decir, contracción, rotación y traslación. De paso, observamos que la razón de áreas de los paralelogramos construidos sobre los vectores unitarios es, simplemente, el determinante de la matriz

$$\begin{bmatrix} k\cos\phi & -l\operatorname{sen}\phi \\ k\operatorname{sen}\phi & l\cos\phi \end{bmatrix}$$

que es kl (¿Recuerda el lector la fórmula del área de un paralelogramo en forma de determinantes descrita en el ejemplo 27 de la lección 2?).

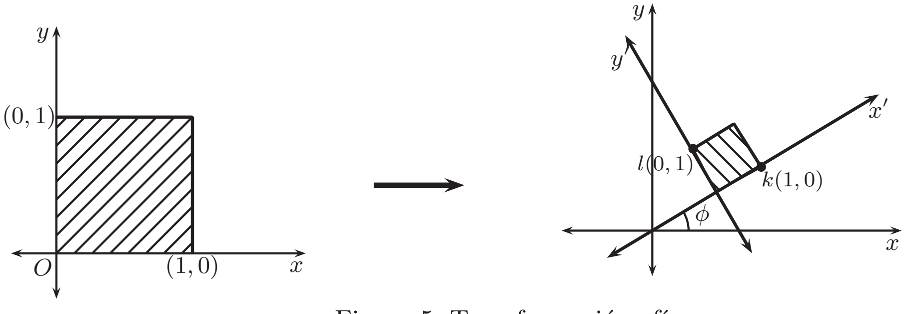


Figura 5. Transformación afín

De otro lado, cabe anotar aquí que un sistema de ecuaciones lineales como

$$\begin{aligned} a_{11}x + a_{12}y &= b_1 \\ a_{21}x + a_{22}y &= b_2 \end{aligned}$$

con

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0$$

y solución $(\, x_s, y_s \,)$ en el sistema cartesiano Oxy, puede verse como el sistema lineal homogéneo

$$\begin{aligned} a_{11}x' + a_{12}y' &= 0 \\ a_{21}x' + a_{22}y' &= 0 \end{aligned}$$

en el sistema cartesiano $Ox'y'$, donde

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix} + \begin{bmatrix} x_s \\ y_s \end{bmatrix}$$

es la transformación afín aplicada al plano. Es decir, una solución del sistema lineal original siempre tiene la forma de una suma entre una solución del sistema homogéneo correspondiente y una solución particular, algo que ya sabíamos de la lección 2.

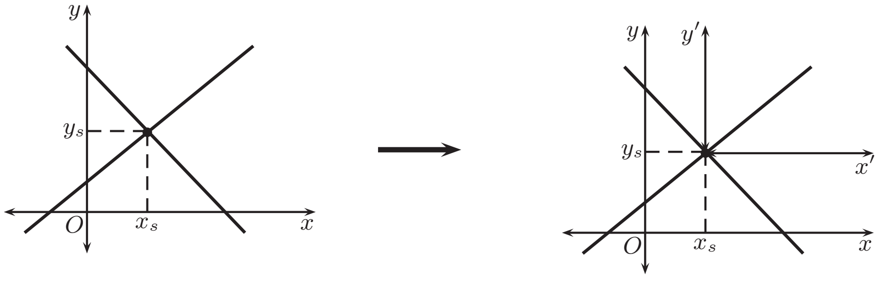


Figura 6. Traslación del origen

1. Transformaciones lineales

La pregunta sobre comportamientos de *invarianza* por transformaciones afines hizo de la geometría analítica una rama aún más indispensable para la investigación en física y en otras ciencias naturales; en particular, dio a la "geometría de invariantes", hoy conocida como *geometría proyectiva*, un impulso esencial. Sobre las consecuencias de estos inmensos desarrollos e ideas de la geometría analítica de los siglos XVII, XVIII y XIX descritos anteriormente, se han basado los estudios de las transformaciones lineales en el submarco específico (y abstracto) del álgebra lineal que han servido bien a los propósitos del análisis y de la física matemática.

Definición 1. (Transformación lineal (Peano (1888)))

Sean V y W espacios vectoriales sobre $\mathbb{R}$;

a) Una *transformación lineal* de V en W es una función $T : V \longrightarrow W$ tal que para todo x, $y \in V$ y $k \in \mathbb{R}$ se tiene que

   i) $T(x + y) = T(x) + T(y)$;

   ii) $T(kx) = kT(x)$;

   es decir, T *"preserva la linealidad"* del espacio V en el espacio W (figura 7).

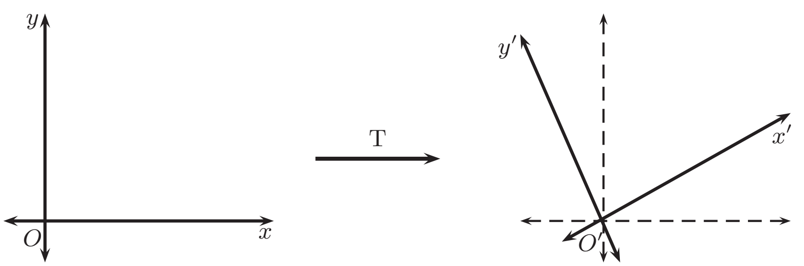


Figura 7. Transformación lineal

b) Si $V = W$, entonces T se llama un *operador lineal* sobre V.

c) T es una *transformación afín* de V en W si $T = L + w$, donde L es una transformación lineal y $w \in W$ es un vector fijo.

Nota 1. (T envía el origen en sí mismo)
Observemos que si T es una transformación lineal de V en W, entonces $T(0) = 0$ (es decir, transforma el $0 \in V$ en el $0 \in W$), pues $T(0) = T(0+0) = T(0) + T(0)$, y así $T(0) = 0$.

Ejemplo 1.
Las transformaciones triviales

$$\begin{aligned} T_1 : V &\longrightarrow V \\ x &\longrightarrow T(x) = x \end{aligned} \qquad\qquad \begin{aligned} T_2 : V &\longrightarrow W \\ x &\longrightarrow T(x) = 0 \end{aligned}$$

son, obviamente, transformaciones lineales.

Ejemplo 2.

a) Las únicas transformaciones lineales $T : \mathbb{R} \longrightarrow \mathbb{R}$ son de la forma

$$T(x) = mx, \quad m \in \mathbb{R} \text{ fijo}$$

pues:

i) Si $x, y \in \mathbb{R}$ y $k \in \mathbb{R}$,

$$T(x+y) = m(x+y) = mx + my = T(x) + T(y)$$
$$T(kx) = m(kx) = k(mx) = kT(x)$$

ii) Si definimos $m \equiv T(1)$, entonces $T(x) = xT(1) = mx$

b) Las únicas transformaciones lineales $T : \mathbb{R}^2 \longrightarrow \mathbb{R}$ son de la forma

$$T(x,y) = ax + by = \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad a, b \in \mathbb{R} \text{ fijos}$$

En efecto:

i) Si $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$, $k \in \mathbb{R}$,

$$\begin{aligned} T((x_1,y_1) + (x_2,y_2)) &= T(x_1 + x_2, y_1 + y_2) \\ &= a(x_1 + x_2) + b(y_1 + y_2) \\ &= [ax_1 + by_1] + [ax_2 + by_2] \\ &= T(x_1,y_1) + T(x_2,y_2) \\ T(k(x_1,y_1)) &= T(kx_1, ky_1) = a(kx_1) + b(ky_1) \\ &= k[ax_1 + by_1] = kT(x_1,y_1) \end{aligned}$$

ii) Si definimos $T(1,0) \equiv a$ y $T(0,1) \equiv b$, entonces

$$T(x,y) = x\,T(1,0) + y\,T(0,1) = ax + by$$

c) De manera similar, las únicas transformaciones lineales $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ son de la forma

$$T(x,y) = (ax+by, cx+dy) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad a,b,c,d \in \mathbb{R}$$

pues si $T(1,0) = (a,c)$ y $T(0,1) = (b,d)$, entonces

$$T(x,y) = x\,T(1,0) + y\,T(0,1) = x\,(a,c) + y\,(b,d) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Ejemplo 3.
En general, si A es una matriz $m \times n$ fija y

$$\begin{aligned} T : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ x &\longrightarrow T(x) = Ax \end{aligned}$$

entonces, claramente, esta T es una transformación lineal:

a) $T(x+y) = A(x+y) = Ax + Ay = T(x) + T(y)$ para todo $x, y \in \mathbb{R}^n$.

b) $T(kx) = A(kx) = k(Ax) = kT(x)$ para todo $x \in \mathbb{R}^n$, $k \in \mathbb{R}$.

Ejemplo 4.
Sea $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ definida por la ecuación $T(x,y) = (x+y, x-y, y)$. Veamos que T es, efectivamente, una transformación lineal.

Solución

Es suficiente observar que $T(x,y) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ y aplicar lo aprendido en el ejemplo 3 con $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}$.

Ejemplo 5.
La transformación que toma cualquier vector $(x,y) \in \mathbb{R}^2$ a través de la línea de 45° para alcanzar su imagen espectral (y,x) puede expresarse como la transformación lineal

$$T(x,y) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Pero si rotamos el plano de tal forma que uno de los ejes nuevos corresponda, precisamente, a la línea de 45º, y el otro eje sea perpendicular a esta línea (figura 8), tendremos que para expresar a (x, y) en términos de $u = (1, 1)$ (línea de 45º) y $v = (-1, 1)$ (línea perpendicular), primero debemos observar que $T(1, 1) = (1, 1)$ y $T(-1, 1) = (1, -1)$ y, por tanto, la nueva descripción es

$$T(x', y') = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = (x' + y', x' - y')$$

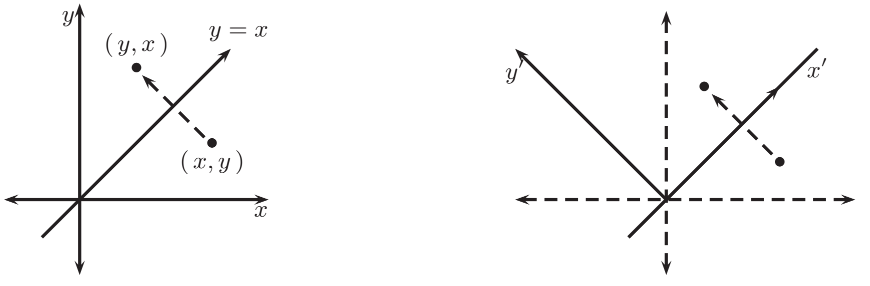


Figura 8

Ejemplo 6. (La trasposición es una transformación lineal)
Sea $T : \mathfrak{M}_{m\times n} \longrightarrow \mathfrak{M}_{n\times m}$ la función definida por

$$T(A) = A^T$$

Veamos que T es una transformación lineal.

Solución

a) Sean A y B dos matrices $m \times n$; entonces

$$T(A + B) = (A + B)^T = A^T + B^T = T(A) + T(B)$$

b) Si tomamos un escalar $k \in \mathbb{R}$ arbitrario, entonces

$$T(kA) = (kA)^T = kA^T = kT(A)$$

Ejemplo 7. (La traza es una transformación lineal)
Sea $T : \mathfrak{M}_{n\times n} \longrightarrow \mathbb{R}$ la función definida por $T(A) = \text{Traza}(A)$. Veamos que T es una transformación lineal.

Solución

a) Sean A y B dos matrices $n \times n$; entonces

$$T(A + B) = \text{Traza}(A + B) = \text{Traza}A + \text{Traza}B = T(A) + T(B)$$

b) Si tomamos un escalar $k \in \mathbb{R}$ cualquiera, entonces

$$T(kA) = \text{Traza}(kA) = k\,\text{Traza}A = kT(A)$$

Ejemplo 8.
Para $n \in \mathbb{N}$ sea $T : P_n \longrightarrow P_{n-1}$ la función definida por

$$T(a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1$$

Mostremos que T es una transformación lineal. Esta podría llamarse la *"función derivada"* por razones que entenderemos en el Volumen 2 (Cálculo).

Solución

a) Sean $p_1(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ y $p_2(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0$ dos polinomios cualquiera en P_n. Entonces

$$\begin{aligned} T(p_1(x) + p_2(x)) &= T((a_n + b_n)x^n + (a_{n-1} + b_{n-1})x^{n-1} + \cdots + \\ &\quad (a_1 + b_1)x + (a_0 + b_0)) \\ &= (a_n + b_n)nx^{n-1} + (a_{n-1} + b_{n-1})(n-1)x^{n-2} + \\ &\quad \cdots + (a_1 + b_1) \\ &= a_n nx^{n-1} + a_{n-1}(n-1)x^{n-2} + \cdots + a_1 + \\ &\quad b_n nx^{n-1} + b_{n-1}(n-1)x^{n-2} + \cdots + b_1 \\ &= T(p_1(x)) + T(p_2(x)) \end{aligned}$$

b) Si tomamos un escalar $k \in \mathbb{R}$ arbitrario, entonces

$$\begin{aligned} T(kp_1(x)) &= T(ka_n x^n + ka_{n-1}x^{n-1} + \cdots + ka_1 x + ka_0) \\ &= ka_n nx^{n-1} + ka_{n-1}(n-1)x^{n-2} + \cdots + ka_1 \\ &= kT(p_1(x)) \end{aligned}$$

Ejemplo 9.
Es fácil mostrar, de manera similar, que $T : P_n \longrightarrow P_{n+1}$ definida por

$$T(a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0) = \frac{a_n}{n+1}x^{n+1} + \frac{a_{n-1}}{n}x^n + \cdots + \frac{a_1}{2}x^2 + a_0 x$$

es una transformación lineal. Esta podría llamarse la *"función integral"* por razones que entenderemos en el Volumen 2 (Cálculo).

Ejemplo 10.
Sea $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ la transformación lineal definida por $T(x, y) = (2x, 2y)$. Describamos la imagen bajo T de los puntos del círculo $x^2 + y^2 = 1$.

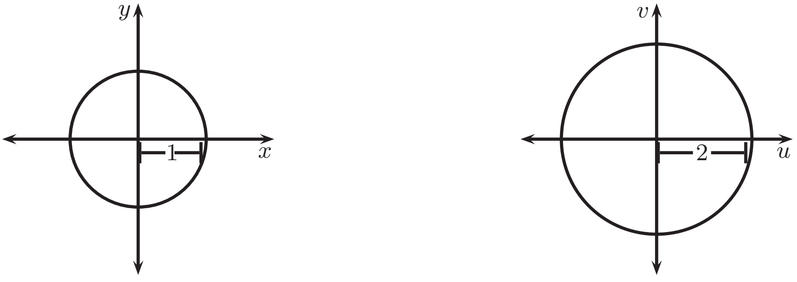


Figura 9

Solución

Sea (x, y) un punto del círculo con centro $(0, 0)$ y radio 1; y sea también $u = 2x,\ v = 2y$. Entonces u, v satisfacen la relación

$$\left(\frac{u}{2}\right)^2 + \left(\frac{v}{2}\right)^2 = 1$$

o, en otras palabras,

$$u^2 + v^2 = 4$$

que muestra que (u, v) está en el círculo de radio 2. De aquí que el círculo de radio 1 se transforma en el de radio 2, ambos con centro en $(0, 0)$ (figura 9). ¿Qué sucedería si la transformación fuera $T(x, y) = (2x, 3y)$?

Ejemplo 11. (Sobre la constancia de la velocidad de la luz: derivación de la transformación lineal de Lorentz)
Muy al final del siglo XIX, se llegó a una contradicción fundamental en física. El bien conocido experimento de Albert A. Michelson [1852–1935], en el cual se medía la velocidad de la luz (aproximadamente 300,000 $\frac{\text{km}}{\text{seg}}$) en dirección del movimiento de la Tierra a lo largo de su órbita alrededor del sol (la velocidad de la Tierra es de aproximadamente 30 $\frac{\text{km}}{\text{seg}}$) mostró irrefutablemente que todos los cuerpos, aún en el vacío, se contraían en la dirección del movimiento. La teoría de esta contracción fue investigada por el físico holandés, Hendrik Lorentz [1853–1928]. Éste mostraba que la contracción es más grande cuando la velocidad del cuerpo se acerca a la velocidad de la luz en el vacío; y que a una velocidad igual a la de la luz, la contracción se hacía infinita. Lorentz derivó las fórmulas de esta contracción.

Poco después, Albert Einstein [1879–1955] introdujo a este problema un punto de vista completamente diferente. Decía que si asumíamos que la propagación de la luz, como cualquier cuerpo ordinario, se realiza mediante la *ley de composición de velocidad* de Galileo, entonces la velocidad de la luz es $c' = c + v$, donde v es la velocidad del observador que se mueve hacia el origen de la luz, y c es la velocidad de la luz para un observador estacionario. Del experimento de Michelson se seguiría que $c' = c$. De hecho, la ley $c' = c + v$ está basada en la transformación lineal

$$\begin{aligned} x' &= x + v_x t \\ t' &= t \end{aligned} \qquad \text{o} \qquad \begin{bmatrix} x' \\ t' \end{bmatrix} = \begin{bmatrix} 1 & v_x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \tag{1}$$

que conecta la coordenada x de un punto relativo a un sistema coordenado I con su coordenada x' relativa al sistema II que tiene su eje x paralelo al eje x del sistema I y se mueve paralelo al eje Ox con velocidad v_x relativo al sistema I (figura 10). Éstas son, precisamente, las fórmulas que Einstein decía que debían cambiarse. Para mostrar la versión más simple posible de lo que éste proponía, supongamos que el objeto se mueve a la velocidad de la luz a lo largo del eje Ox. Entonces las nuevas coordenadas son de la forma

$$\begin{aligned} x' &= a_1 x + d_1 t \\ t' &= a_2 x + d_2 t \end{aligned} \qquad \text{o} \qquad \begin{bmatrix} x' \\ t' \end{bmatrix} = \begin{bmatrix} a_1 & d_1 \\ a_2 & d_2 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \tag{2}$$

donde a_1, a_2, d_1, d_2 son ciertas constantes. Einstein demuestra que $a_1 = d_2 = \dfrac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}$, $d_1 = \dfrac{-v}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}$ y $a_2 = \dfrac{\frac{-v}{c^2}}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}$, donde $v = \left(\dfrac{\tau^2 - 1}{\tau^2 + 1}\right) c$ y τ es el "coeficiente de contracción-expansión" en los ejes x, t de la transformación lineal dada por (2).

Así, $x' = \dfrac{x - vt}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}$ y $t' = \dfrac{t - \frac{vx}{c^2}}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}$, que son las fórmulas de Einstein-Lorentz. Obsérvese ahora que, *sólo si v es muy pequeña con respecto a c*, entonces obtenemos que $x' = x + v_x t$ y $t' = t$ que son las fórmulas originales de Michelson del sistema (1).

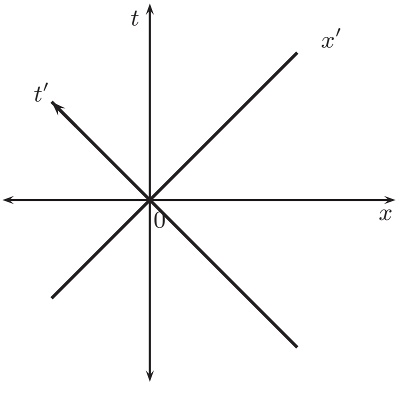


Figura 10

a. Transformaciones ortogonales

Una clase muy especial de transformaciones lineales son aquellas $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ que satisfacen la condición de *preservar las longitudes de los vectores*; es decir, que $||\, T(\, x\,)\, || = ||\, x\, ||$ para todo $x \in \mathbb{R}^n$. Casos de este tipo de transformaciones ya han surgido: por ejemplo, si T es una “rotación de un ángulo ϕ”, es decir, si

$$T(\, x, y\,) = \begin{bmatrix} \cos\,\phi & -\operatorname{sen}\,\phi \\ \operatorname{sen}\,\phi & \cos\,\phi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

tenemos que

$$\begin{aligned} ||\, T(\, x, y\,)\, ||^2 &= ||\, (\, x\cos\,\phi - y\operatorname{sen}\,\phi, x\operatorname{sen}\,\phi + y\cos\,\phi\,)\, ||^2 \\ &= [\, x^2\cos^2\phi - 2xy\cos\,\phi\operatorname{sen}\,\phi + y^2\operatorname{sen}^2\phi\,] + \\ &\quad\; [\, x^2\operatorname{sen}^2\phi + 2xy\operatorname{sen}\,\phi\cos\,\phi + y^2\cos^2\phi\,] \\ &= x^2(\, \cos^2\phi + \operatorname{sen}^2\phi\,) + y^2(\, \operatorname{sen}^2\phi + \cos^2\phi\,) = x^2 + y^2 \\ &= ||\, (\, x, y\,)\, ||^2 \end{aligned}$$

Movimientos rígidos del plano sobre sí mismo, o este movimiento más una rotación en un ángulo ϕ, son transformaciones ortogonales del plano; y movimientos rígidos del espacio tales como reflexiones alrededor de planos se llaman *transformaciones ortogonales* del espacio. Otros ejemplos de este tipo de movimiento son

$$\begin{aligned} T : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (\, x, y\,) &\longrightarrow (\, y, x\,) \end{aligned} \qquad\qquad \begin{aligned} T : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (\, x, y, z\,) &\longrightarrow (\, z, -y, x\,) \end{aligned}$$

En general, tenemos la siguiente definición:

Definición 2. (Transformación ortogonal)
Diremos que una transformación lineal $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ es *ortogonal* si, y sólo si, para todo $x \in \mathbb{R}^n$,

$$||\, T(\, x\,)\, || = ||\, x\, ||$$

Y podría no ser sorprendente que existiera relación entre las transformaciones ortogonales que acabamos de definir, y las matrices ortogonales que introdujimos al final de la lección 5 anterior. Y en efecto es así:

Teorema 1.

Sea $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ una transformación lineal que está definida por $T(\, x\,) = Ax$ para alguna matriz $A \in \mathfrak{M}_{n\times n}$; entonces T es ortogonal si, y sólo si, A es ortogonal (es decir, si $A^T = A^{-1}$).

Demostración

a) Si A es ortogonal, entonces

$$||\,T(\,x\,)\,||^2 = [T(\,x\,)]^T[T(\,x\,)] = (\,Ax\,)^T(\,Ax\,) = x^TA^TAx = x^Tx = ||\,x\,||^2$$

b) Si $||\,T(\,x\,)\,|| = ||\,x\,||$ para todo $x \in \mathbb{R}^n$, entonces, de manera similar a lo obtenido en a),

$$x^TA^TAx = x^Tx$$

Escribamos $A^TA = (\,a_{ij}\,)$. Entonces para $x = e_i = (\,0, 0, \ldots, 1, \ldots, 0\,)$, con el 1 en la posición $i = 1, 2, \ldots, n$, se tiene que

$$a_{ij} = \begin{cases} 1 & \text{si} \quad j = i \\ 0 & \text{si} \quad j \neq i \end{cases}$$

Así, $A^TA = I_n$ y, por tanto, A es ortogonal. ■

Ejemplo 12.

Probemos, utilizando el teorema anterior, que $T(\,x\,) = Ax$ definida con

$$A = \begin{bmatrix} \frac{6}{7} & \frac{2}{7} & \frac{3}{7} \\ \frac{2}{7} & \frac{3}{7} & -\frac{6}{7} \\ -\frac{3}{7} & \frac{6}{7} & \frac{2}{7} \end{bmatrix}$$

es una transformación ortogonal de $\mathbb{R}^3$ en $\mathbb{R}^3$.

Solución

La transformación T es ortogonal, ya que A es una matriz ortogonal:

$$AA^T = \begin{bmatrix} \frac{6}{7} & \frac{2}{7} & \frac{3}{7} \\ \frac{2}{7} & \frac{3}{7} & -\frac{6}{7} \\ -\frac{3}{7} & \frac{6}{7} & \frac{2}{7} \end{bmatrix} \begin{bmatrix} \frac{6}{7} & \frac{2}{7} & -\frac{3}{7} \\ \frac{2}{7} & \frac{3}{7} & \frac{6}{7} \\ \frac{3}{7} & -\frac{6}{7} & \frac{2}{7} \end{bmatrix} = \begin{bmatrix} \frac{36+4+9}{49} & \frac{12+6-18}{49} & \frac{-18+12+6}{49} \\ \frac{12+6-18}{49} & \frac{4+9+36}{49} & \frac{-6+18-12}{49} \\ \frac{-18+12+6}{49} & \frac{-6+18-12}{49} & \frac{9+36+4}{49} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Ejemplo 13.

Mostremos, utilizando el teorema 1, que $T(x) = Ax$ con

$$A = \begin{bmatrix} \frac{7}{9} & -\frac{4}{9} & \frac{4}{9} \\ -\frac{4}{9} & \frac{1}{9} & \frac{8}{9} \\ \frac{4}{9} & \frac{8}{9} & \frac{1}{9} \end{bmatrix}$$

es una transformación ortogonal de $\mathbb{R}^3$ en $\mathbb{R}^3$.

Solución

La transformación T es ortogonal, ya que A es una matriz ortogonal:

$$AA^T = \begin{bmatrix} \frac{7}{9} & -\frac{4}{9} & \frac{4}{9} \\ -\frac{4}{9} & \frac{1}{9} & \frac{8}{9} \\ \frac{4}{9} & \frac{8}{9} & \frac{1}{9} \end{bmatrix} \begin{bmatrix} \frac{7}{9} & -\frac{4}{9} & \frac{4}{9} \\ -\frac{4}{9} & \frac{1}{9} & \frac{8}{9} \\ \frac{4}{9} & \frac{8}{9} & \frac{1}{9} \end{bmatrix} = \begin{bmatrix} \frac{49+16+16}{81} & \frac{-28-4+32}{81} & \frac{28-32+4}{81} \\ \frac{-28-4+32}{81} & \frac{16+1+64}{81} & \frac{-16+8+8}{81} \\ \frac{28-32+4}{81} & \frac{-16+8+8}{81} & \frac{16+1+64}{81} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Ejercicios 1

1) ¿Cuáles de las siguientes funciones $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ son transformaciones lineales?:

   a) $T(x, y) = (3 + x, y)$
   b) $T(x, y) = (x, y - 5)$
   c) $T(x, y) = (y, x)$
   d) $T(x, y) = (x^2, y)$
   e) $T(x, y) = (\operatorname{sen} x, \cos x)$
   f) $T(x, y) = (xy, y)$

2) ¿Existe una transformación lineal $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ tal que $T(1, -1, 1) = (2, 0)$ y $T(1, 1, 1) = (0, 3)$?

3) Pruebe que $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por

$$T(x, y, z) = (x - y + z, 3x + y, -x - y - 5z)$$

   es una transformación lineal. Escríbala en la forma

$$T(x, y, z) = A \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

   para cierta matriz 3×3, A, por identificar.

*4) Interprete *geométricamente* cada una de las transformaciones lineales $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ o $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ que han aparecido en esta sección 1.

5) Geométricamente ¿qué significado tiene una transformación $T(\,x\,) = Ax$ si A es una matriz *simétrica*?

6) Si $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ es definida por $T(\,x, y\,) = \left(\dfrac{x}{3}, \dfrac{y}{3}\right)$ ¿Cuál es la imagen bajo T de la elipse $\dfrac{x^2}{9} + \dfrac{y^2}{16} = 1$?

7) Describa la imagen en $\mathbb{R}^2$ de la recta $P + tA$ bajo la transformación lineal T, donde:

a) $T(\,x, y\,) = (\,2x + y, y\,)$
b) $T(\,x, y\,) = (\,y, x\,)$
c) $T(\,x, y\,) = (\,5x, y - x\,)$

8) Describa la imagen del plano $(X - P)N = 0$ en $\mathbb{R}^3$ bajo la transformación lineal T, donde:

a) $T(\,x, y, z\,) = (\,3x - 2y + z, 2x - y + 2z, x - 3z\,)$
b) $T(\,x, y, z\,) = (\,-x + 4y + 2z, x - y - z, y + 4z\,)$
c) $T(\,x, y, z\,) = (\,2x + 2y + 2z, x - y, y + z\,)$

9) Sea $T(\,x\,) = Ax$, donde:

a) $A = \begin{bmatrix} 6 & 8 \\ 8 & -6 \end{bmatrix}$ b) $A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$

c) $A = \begin{bmatrix} 0.50 & 0.25 & 0.25 \\ 0.25 & 0.50 & 0.25 \\ 0.25 & 0.25 & 0.50 \end{bmatrix}$ d) $A = \begin{bmatrix} 0 & 18 & -24 \\ -18 & 0 & 40 \\ 24 & -40 & 0 \end{bmatrix}$

e) $A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$ f) $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

¿En qué casos es T una transformación lineal ortogonal?

10) Encuentre la tercera columna de la matriz

$$A = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & ? \\ \frac{1}{\sqrt{3}} & 0 & ? \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & ? \end{bmatrix}$$

de tal manera que sea ortogonal.

11) Muestre, utilizando la definición $||\,T(\,x\,)\,|| = ||\,x\,||$, que $T(\,x\,) = Ax$ con

$$A = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

es una transformación ortogonal de $\mathbb{R}^3$ en $\mathbb{R}^3$.

2. Núcleo e imagen: dos subespacios asociados a una transformación lineal

A diferencia de las funciones no-lineales, uno de los elementos claves en el comportamiento global de las transformaciones lineales es que pueden describirse muy simplemente, como iremos viendo en las secciones que siguen. Dos de estas claves son precisamente los (dos) siguientes subespacios que surgen naturalmente de la estructura lineal de la transformación.

Definición 3. (Núcleo e imagen)
Sea la función $T : V \longrightarrow W$ una transformación lineal cualquiera.

a) Definimos el *núcleo* (*kernel* o *espacio nulo*) de T así:

$$\text{Nu}(\,T\,) = \{\,x \in V \,/\, T(\,x\,) = 0\,\}$$

b) Definimos la *imagen* de T así:

$$\text{Im}(\,T\,) = \{\,y \in W \,/\ \text{existe}\, x \in V \ \text{y}\, T(\,x\,) = y\,\}$$

(es decir, $\text{Im}(\,T\,)$ es el rango de la función T)

Teorema 2.
Nu$(\,T\,)$ y Im$(\,T\,)$ son subespacios de V y W, respectivamente.

Demostración

a) Ver que el núcleo de T, $\text{Nu}(\,T\,)$, es un subespacio de V es fácil:

i) Si x, $y \in \text{Nu}(\,T\,)$, entonces $T(\,x\,) = 0$ y $T(\,y\,) = 0$; luego

$$T(\,x + y\,) = T(\,x\,) + T(\,y\,) = 0 + 0 = 0$$

y así $x + y \in \text{Nu}(\,T\,)$.

   ii) Si $k \in \mathbb{R}$, $x \in \text{Nu}(T)$, entonces $T(kx) = kT(x) = k \cdot 0 = 0$.

b) De manera similar, ver que la imagen de T, $\text{Im}(T)$, es un subespacio de W es también fácil:

   i) Si z, $w \in \text{Im}(T)$, entonces existen x, $y \in V$ tales que $T(x) = z$ y $T(y) = w$. Luego, $T(x + y) = T(x) + T(y) = z + w$; así $z + w \in \text{Im}(T)$.

   ii) Si $k \in \mathbb{R}$, $z \in \text{Im}(T)$, entonces, puesto que existe $x \in V$ tal que $T(x) = z$, se tendrá que $T(kx) = kT(x) = kz$; luego $kz \in \text{Im}(T)$. ■

Si V es finito-dimensional, a la dimensión de $\text{Im}(T)$ la llamaremos el *rango de* T y la denotaremos por $\rho(T)$. De manera similar, si V es finito-dimensional, a la dimensión de $\text{Nu}(T)$ la llamaremos la *nulidad de* T y la denotaremos por $\text{nul}(T)$.

Teorema 3. ***(Teorema fundamental del álgebra lineal (Grassmann (1846)))***
Supongamos que V es finito-dimensional. Si $T : V \longrightarrow W$ es una transformación lineal, entonces

$$\dim V = \rho(T) + nul(T)$$

Demostración

Sea $\dim V = n$, y $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ una base para el núcleo de T con $m \leq n$ y completemos esta base a una de V: $\{\alpha_1, \alpha_2, \ldots, \alpha_m, \alpha_{m+1}, \ldots, \alpha_n\}$. Es suficiente probar que $\{T(\alpha_{m+1}), \ldots, T(\alpha_n)\}$ conforman una base para $\text{Im}(T)$. En efecto,

a) $\{T(\alpha_{m+1}), \ldots, T(\alpha_n)\}$ generan a $\text{Im}(T)$, pues si $T(x) \in \text{Im}(T)$ para algún $x \in V$, entonces

$$x = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m + k_{m+1}\alpha_{m+1} + \cdots + k_n\alpha_n$$

para algunos escalares k_i's, $i = 1, 2, \ldots, n$; así

$$T(x) = k_{m+1}T(\alpha_{m+1}) + \cdots + k_nT(\alpha_n)$$

pues $\alpha_1, \alpha_2, \ldots, \alpha_m \in \text{Nu}(T)$.

b) $\{T(\alpha_{m+1}), \ldots, T(\alpha_n)\}$ son linealmente independientes se ve cuando asumimos que existen $k_{m+1}, \ldots, k_n$ tales que

$$k_{m+1}T(\alpha_{m+1}) + \cdots + k_nT(\alpha_n) = 0$$

entonces $T(k_{m+1}\alpha_{m+1}+\cdots+k_n\alpha_n)=0$ y así, $k_{m+1}\alpha_{m+1}+\cdots+k_n\alpha_n \in \mathrm{Nu}(T)$ y como $\{\alpha_1,\alpha_2,\ldots,\alpha_m\}$ es una base para $\mathrm{Nu}(T)$, existen, a su vez, escalares $k_1,\ldots,k_n$ tales que

$$k_1\alpha_1+\ldots+k_m\alpha_m = k_{m+1}\alpha_{m+1}+\cdots+k_n\alpha_n$$

o, de otra forma,

$$k_1\alpha_1+\ldots+k_m\alpha_m - k_{m+1}\alpha_{m+1}-\cdots-k_n\alpha_n=0$$

Pero puesto que $\{\alpha_1,\ldots,\alpha_m,\ldots,\alpha_n\}$ es una base para V, entonces serán linealmente independientes y, de esta manera, $k_1=\cdots=k_m=k_{m+1}=\cdots=k_n=0$, lo que finaliza la prueba. ■

Ejemplo 14. (Más luz sobre las soluciones a los sistemas lineales homogéneos)
Sea A una matriz $m\times n$ y $T:\mathbb{R}^n \longrightarrow \mathbb{R}^m$ definida por $T(x)=Ax$. Entonces su núcleo

$$\mathrm{Nu}(T)=\{x\in\mathbb{R}^n\,/\,Ax=0\}$$

es el conjunto de *soluciones* al sistema homogéneo $Ax=0$.

De otro lado, por definición,

$$\mathrm{Im}(T)=\{y\in\mathbb{R}^m\,/\,Ax=y,\quad \text{para algún } x\in\mathbb{R}^n\}$$

y si escribimos la matriz A de la forma

$$A=[\,A_1\,|\,A_2\,|\,\ldots\,|\,A_n\,]$$

donde A_i es la columna i de la matriz A, entonces

$$\begin{aligned}\mathrm{Im}(T)&=\{y\in\mathbb{R}^m\,/\,y=x_1A_1+x_2A_2+\cdots+x_nA_n \ \text{ para algún}\\ &\qquad (x_1,x_2,\ldots,x_n)\in\mathbb{R}^n\}\\ &=\langle A_1,A_2,\ldots,A_n\rangle;\end{aligned}$$

es decir, la imagen de T es el *espacio generado por las columnas de* A, y así el rango de T es la dimensión de este último espacio.

Ahora: de acuerdo con el teorema 3, se tiene que

$$n=\dim\mathbb{R}^n=\rho(T)+\mathrm{nul}(T)$$

De allí, $\mathrm{nul}(T)=0$ si, y sólo si, $\rho(T)=n$; y así, tendremos que el sistema lineal homogéneo $Ax=0$ tiene *sólo la solución trivial* $x=0$ si, y sólo si, las n columnas de A generan $\mathbb{R}^n$; y esto, si y sólo si, las columnas de A son linealmente independientes. Luego, en particular, si $\rho(T)<n$, entonces $Ax=0$ tiene *infinitas soluciones.*[1]

[1] ¿Y qué sucede con el sistema lineal $AX=b$? es decir, ¿cuántas soluciones tendrá en estos casos este sistema?

Ejemplo 15.
Para las siguientes transformaciones lineales:

a) $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por $T(x, y) = (x + 2y, -y)$

b) $T : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ definida por $T(x, y, z, w) = (x + w, y - z)$

encontremos $\mathrm{Im}(T)$ y $\mathrm{Nu}(T)$, sus dimensiones, y corroboremos el teorema fundamental del álgebra lineal.

Solución

a)
$$\begin{aligned}\mathrm{Im}(T) &= \{(x', y') \in \mathbb{R}^2 \,/\, \text{existe } (x, y) \in \mathbb{R}^2 \text{ tal que } T(x, y) \\ &\qquad = (x', y')\} \\ &= \{(x', y') \in \mathbb{R}^2 \,/\, \text{existe } (x, y) \in \mathbb{R}^2 \text{ tal que } (x + 2y, -y) \\ &\qquad = (x', y')\} \\ &= \langle (1, 0), (2, -1) \rangle \\ &= \mathbb{R}^2 \\ \mathrm{Nu}(T) &= \{(x, y) \in \mathbb{R}^2 \,/\, T(x, y) = 0\} \\ &= \{(x, y) \in \mathbb{R}^2 \,/\, (x + 2y, -y) = 0\} \\ &= \{(0, 0)\}\end{aligned}$$

En este caso, $\dim V = 2$, $\rho(T) = 2$ y $\mathrm{nul}(T) = 0$. Luego, $\dim V = \rho(T) + \mathrm{nul}(T)$.

b)
$$\begin{aligned}\mathrm{Im}(T) &= \{(x', y') \in \mathbb{R}^2 \,/\, \text{existe } (x, y, z, w) \in \mathbb{R}^4 \text{ tal que } T(x, y, z, w) \\ &\qquad = (x', y')\} \\ &= \{(x', y') \in \mathbb{R}^2 \,/\, \text{existe } (x, y, z, w) \in \mathbb{R}^4 \text{ tal que } (x + w, y - z) \\ &\qquad = (x', y')\}\end{aligned}$$

$$\begin{aligned}&= \langle (1, 0), (0, 1), (0, -1), (1, 0) \rangle \\ &= \langle (1, 0), (0, 1) \rangle \\ &= \mathbb{R}^2 \\ \mathrm{Nu}(T) &= \{(x, y, z, w) \in \mathbb{R}^4 \,/\, T(x, y, z, w) = (0, 0, 0, 0)\} \\ &= \{(x, y, z, w) \in \mathbb{R}^4 \,/\, (x + w, y - z) = (0, 0)\} \\ &= \{(x, y, z, w) \in \mathbb{R}^4 \,/\, x = -w, y = z\} \\ &= \langle (1, 0, 0, -1), (0, 1, 1, 0) \rangle\end{aligned}$$

Observemos que $\dim V = 4$, $\rho(T) = 2$ y $\mathrm{nul}(T) = 2$. Por tanto, $\dim V = \rho(T) + \mathrm{nul}(T)$.

Y las definiciones anteriores nos permiten ahora caracterizar las transformaciones lineales biyectivas:

Teorema 4. ***(Biyectividad y linealidad)***
Si $T : V \longrightarrow V$ es una transformación lineal con $\dim V = n$, entonces las cuatro siguientes afirmaciones son equivalentes:

a) *T es biyectiva.*

b) *$nul(T) = 0$*

c) *$\rho(T) = n$*

d) *Si $\{\alpha_1, \ldots, \alpha_n\}$ es una base de V, entonces $\{T(\alpha_1), \ldots, T(\alpha_n)\}$ es también una base para V.*

Demostración

a) (Supongamos a) y probemos b)) Si T es biyectiva, entonces no puede existir $x \neq 0$ tal que $T(x) = 0$; y así $\mathrm{Nu}(T) = \{0\}$ y, por tanto, $\mathrm{nul}(T) = 0$.

b) (Supongamos b) y probemos c)) Si $\mathrm{nul}(T) = 0$, entonces, por el teorema fundamental del álgebra lineal, $\rho(T) = n$.

c) (Supongamos c) y probemos d)). Si $\rho(T) = n$, entonces $\mathrm{Nu}(T) = \{0\}$ y, por tanto, si $k_1T(\alpha_1) + \cdots + k_nT(\alpha_n) = 0$ para ciertos $k_1, k_2, \ldots, k_n$, entonces
$$T(k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n) = 0\,,$$
y así $k_1\alpha_1 + \cdots + k_n\alpha_n = 0$; que, a su vez, implica que $k_1 = \cdots = k_n = 0$, pues $\{\alpha_1, \cdots, \alpha_n\}$ es una base de V. Así, $\{T(\alpha_1), \ldots, T(\alpha_n)\}$ es un conjunto linealmente independiente de T y, por tener n vectores, es una base para V.

d) (Supongamos d) y probemos a)). Sea $\{\alpha_1, \ldots, \alpha_n\}$ una base para V. Si $T(x) = T(y)$, entonces, como $x = \sum_{i=1}^{n} k_i\alpha_i$ y $y = \sum_{i=1}^{n} l_i\alpha_i$ para ciertos $k_i, l_i \in \mathbb{R}$ se tendrá que $T(\sum_{i=1}^{n} k_i\alpha_i) = T(\sum_{i=1}^{n} l_i\alpha_i)$ o que $T(\sum_{i=1}^{n}(k_i - l_i)\alpha_i) = 0$, o $\sum_{i=1}^{n}(k_i - l_i)T(\alpha_i) = 0$; pero como $\{T(\alpha_i)\}_{i=1}^{n}$ es una base para V, entonces $k_i = l_i$ para todo $i = 1, \ldots, n$, luego $x = y$. Así hemos probado que T es uno a uno. Ahora: para probar que T es sobre, tomemos $z \in V$ y encontremos $x \in V$ tal que $T(x) = z$. Pero puesto que $z = \sum_{i=1}^{n} m_i T(\alpha_i)$ para algunos $m_i \in \mathbb{R}$, entonces basta tomar $x = \sum_{i=1}^{n} m_i \alpha_i$ y tendremos que $T(x) = z$. ■

Ejemplo 16.
Sea $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ la transformación lineal definida por $T(x, y) = (x + y, x - y)$. Veamos que T es biyectiva.

Solución

De acuerdo con el teorema anterior, es suficiente ver que $\text{Nu}T = \{(0,0)\}$; y esto es así porque $T(x, y) = (0,0)$ cuando, y sólo cuando, $x + y = 0$ y $x - y = 0$; y esto último es cierto cuando, y sólo cuando $x = y = 0$.

Ejemplo 17.
Sea $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ la transformación lineal definida por $T(x, y, z) = (3x + 2y - z, x - y - z, 2x + y - z)$. Veamos que T es biyectiva.

Solución

Por el teorema 4, basta mostrar que $\text{Nu}(T) = \{(0,0,0)\}$. Veamos esto.

$$\begin{aligned}\text{Nu}(T) &= \{(x, y, z) \in \mathbb{R}^3 \,/\, T(x, y, z) = 0\}\\ &= \{(x, y, z) \in \mathbb{R}^3 \,/\, (3x + 2y - z, x - y - z, 2x + y - z) = (0,0,0)\}\end{aligned}$$

es decir, $(x, y, z) \in \text{Nu}(T)$ si, y sólo si,

$$\begin{aligned}3x + 2y - z &= 0\\ x - y - z &= 0\\ 2x + y - z &= 0\end{aligned}$$

La matriz aumentada de este sistema de ecuaciones es

$$\left[\begin{array}{rrr|r} 3 & 2 & -1 & 0\\ 1 & -1 & -1 & 0\\ 2 & 1 & -1 & 0\end{array}\right] \begin{array}{l} F_1\\ F_2\\ F_3\end{array}$$

Por el método de eliminación gaussiana obtenemos que

$$\left[\begin{array}{rrr|r} 1 & \frac{2}{3} & -\frac{1}{3} & 0\\ 1 & -1 & -1 & 0\\ 2 & 1 & -1 & 0\end{array}\right] \quad F_1 \longleftrightarrow \tfrac{1}{3}F_1$$

$$\left[\begin{array}{rrr|r} 1 & \frac{2}{3} & -\frac{1}{3} & 0\\ 0 & \frac{5}{3} & \frac{2}{3} & 0\\ 0 & \frac{1}{3} & \frac{1}{3} & 0\end{array}\right] \quad \begin{array}{l} F_2 \longleftrightarrow F_1 - F_2\\ F_3 \longleftrightarrow 2F_1 - F_3\end{array}$$

$$\begin{bmatrix} 1 & \frac{2}{3} & -\frac{1}{3} & | & 0 \\ 0 & 1 & \frac{2}{5} & | & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & | & 0 \end{bmatrix} \quad F_2 \longleftrightarrow \tfrac{3}{5}F_2$$

$$\begin{bmatrix} 1 & 0 & -\frac{9}{15} & | & 0 \\ 0 & 1 & \frac{2}{5} & | & 0 \\ 0 & 0 & -\frac{1}{5} & | & 0 \end{bmatrix} \quad \begin{matrix} F_1 \longleftrightarrow F_1 - \frac{2}{3}F_2 \\ \\ F_3 \longleftrightarrow \frac{1}{3}F_2 - F_3 \end{matrix}$$

$$\begin{bmatrix} 1 & 0 & \frac{9}{15} & | & 0 \\ 0 & 1 & \frac{2}{5} & | & 0 \\ 0 & 0 & 1 & | & 0 \end{bmatrix} \quad F_3 \longleftrightarrow -5F_3$$

$$\begin{bmatrix} 1 & 0 & 0 & | & 0 \\ 0 & 1 & 0 & | & 0 \\ 0 & 0 & 1 & | & 0 \end{bmatrix} \quad \begin{matrix} F_1 \longleftrightarrow F_1 + \frac{9}{5}F_3 \\ F_2 \longleftrightarrow F_2 - \frac{2}{5}F_3 \end{matrix}$$

Por tanto, la única solución de este sistema de ecuaciones es $x = y = z = 0$. Así, $\mathrm{Nu}(T) = \{(0,0,0)\}$ y T es biyectiva. ▲

Ejercicios 2

1) Sea $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por

$$T(x,y,z) = \begin{bmatrix} 1 & 3 & -1 \\ 2 & -1 & -1 \\ 1 & 4 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

   a) Verifique que T es una transformación lineal.
   b) ¿Cuál es el núcleo de T y la imagen de T?
   c) ¿Cuáles son las dimensiones $\rho(T)$ y nul (T)?
   d) Compruebe el teorema fundamental del álgebra lineal.

2) a) Encuentre el rango de las transformaciones lineales $T(x) = Ax$, donde

   i) $A = \begin{bmatrix} 1 & 1 \\ \frac{1}{3} & \frac{1}{3} \\ 0 & 0 \end{bmatrix}$ ii) $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \end{bmatrix}$

   iii) $A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ iv) $A = \begin{bmatrix} 7 & 1 & 3 & 6 \\ 0 & 3 & 1 & 1 \\ -7 & -1 & -3 & 5 \end{bmatrix}$

b) Confirme que el rango coincide con la dimensión del espacio generado por las columnas de la matriz correspondiente.

$*3$) Sea V un espacio finito-dimensional de dimensión n y $\beta = \{\beta_1, \beta_2, \ldots, \beta_n\}$ una base "ordenada" para V. Dado $x \in V$, sabemos que existe una *única* n-pla $(c_1, c_2, \ldots, c_n)$ de escalares tales que

$$x = \sum_{i=1}^{n} c_i \, \beta_i \tag{1}$$

A c_i lo llamaremos la *coordenada i-ésima de x relativa a la base ordenada* β. Por tanto, podemos asociar a cada vector $x \in V$ un vector $(c_1, c_2, \ldots, c_n) \in \mathbb{R}^n$ que usualmente se denota por $[x]_\beta$.

Ahora supongamos que tenemos otra base ordenada de V: $\gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_n\}$. Muestre que si x está representada en la base β por $[x]_\beta$, y está representada en la base γ por $[x]_\gamma$, entonces $[x]_\gamma = P^{-1}[x]_\beta$ para cierta matriz invertible P conocida como *matriz de paso de la base* β *a la base* γ [Indicación: P tiene como columna j al vector $[\gamma_j]_\beta$].

3. Transformaciones lineales y matrices

Como tal vez se haya entendido a través de las discusiones anteriores, es posible mostrar que toda transformación lineal de un espacio vectorial finito-dimensional V en otro espacio vectorial finito-dimensional W, puede describirse a través de una matriz. Ilustremos, inicialmente con un par de ejemplos, cómo pueden encontrarse estas matrices.

Ejemplo 18.
Consideremos la transformación lineal $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ definida por la ecuación $T(x, y) = (x + y, x - y, y)$. Con el fin de descubrir una matriz asociada a la transformación lineal dada, evaluemos T sobre la base canónica $\{(1, 0), (0, 1)\}$ de $\mathbb{R}^2$. Entonces $T(1, 0) = (1, 1, 0)$ y $T(0, 1) = (1, -1, 1)$. Ahora: observemos que

$$\begin{aligned} T(x, y) &= T(x(1, 0) + y(0, 1)) \\ &= T(x(1, 0)) + T(y(0, 1)) \\ &= xT(1, 0) + yT(0, 1) \\ &= x(1, 1, 0) + y(1, -1, 1) \end{aligned}$$

De la última igualdad podemos obtener una ecuación matricial para T:

$$T(x, y) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Notemos que hemos colocado, sobre las *columnas* de la matriz, los vectores $T(1,0)$ y $T(0,1)$; es decir, hemos colocado las imágenes de los elementos de la base canónica de $\mathbb{R}^2$, $T(1,0)$ y $T(0,1)$.

Ejemplo 19.
Para la transformación lineal $T:\mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por $T(x,y,z) = (2x - y + z, x + 3y - z, 2x + 2y - 2z)$, observemos que

$$\begin{aligned} T(x,y,z) &= T(x(1,0,0) + y(0,1,0) + z(0,0,1)) \\ &= T(x(1,0,0)) + T(y(0,1,0)) + T(z(0,0,1)) \\ &= xT(1,0,0) + yT(0,1,0) + zT(0,0,1) \\ &= x(2,1,2) + y(-1,3,2) + z(1,-1,-2) \end{aligned}$$

o, equivalentemente, que

$$T(x,y,z) = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 3 & -1 \\ 2 & 2 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \blacktriangle$$

Pero lo anterior no es sólo cierto en estos casos particulares de los ejemplos 18 y 19. Es, de hecho, un importante resultado que muestra la profunda *conexión entre las matrices y las transformaciones lineales.* Veamos su formulación formal.

Teorema 5. ***(Transformaciones lineales y matrices (Peano(1888)))***
Si $T:V \longrightarrow W$ es una transformación lineal cualquiera, $\beta = \{\alpha_1, \ldots, \alpha_n\}$ una base para V, y $\beta' = \{\beta_1, \ldots, \beta_m\}$ una base para W, entonces existe una única matriz $A \in \mathfrak{M}_{m\times n}$ tal que para todo $x \in V$,

$$[T(x)]_{\beta'} = A[x]_\beta$$

donde

a) $[T(x)]_{\beta'}$ *es el vector* $m \times 1$ *de coeficientes* $l_1, l_2, \ldots, l_m$ *tales que*

$$T(x) = l_1\beta_1 + l_2\beta_2 + \cdots + l_m\beta_m = \sum_{j=1}^{m} l_j\beta_j \tag{1}$$

b) $[x]_\beta$ *es el vector* $n \times 1$ *de coeficientes* $k_1, k_2, \ldots, k_n$ *tales que*

$$x = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n \tag{2}$$

Demostración

Sea $x \in V$ cualquiera. Si $x = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n = \sum_{i=1}^{n} k_i\alpha_i$, entonces

$$T(x) = k_1T(\alpha_1) + k_2T(\alpha_2) + \cdots + k_nT(\alpha_n) = \sum_{i=1}^{n} k_iT(\alpha_i)$$

Pero para $i = 1, 2, \ldots, n$ existen unos únicos coeficientes $\{a_{ij}\}$ con $j = 1, \ldots, m$ tales que

$$T(\alpha_i) = a_{1i}\beta_1 + a_{2i}\beta_2 + \cdots + a_{mi}\beta_m = \sum_{j=1}^{m} a_{ji}\beta_j \tag{3}$$

Luego

$$T(x) = \sum_{i=1}^{n} k_iT(\alpha_i) = \sum_{i=1}^{n} k_i\left(\sum_{j=1}^{m} a_{ji}\beta_j\right) = \sum_{j=1}^{m}\left(\sum_{i=1}^{n} a_{ji}k_i\right)\beta_j \tag{4}$$

Comparando (1) con (4) definimos

$$l_j \equiv \sum_{i=1}^{n} a_{ji}k_i$$

Luego si $A = [a_{ij}]_{m\times n}$, entonces

$$[T(x)]_{\beta'} = (l_1, \ldots, l_m)^T = \left(\sum_{i=1}^{n} a_{1i}k_i, \ldots, \sum_{i=1}^{n} a_{mi}k_i\right)^T = A[x]_\beta$$

donde $[x]_\beta = (k_1, ..., k_n)^T$. ■

Nota 2.

i) A la matriz A del teorema anterior se le llama la *matriz de la transformación lineal T relativa a las bases β* y β', y la notaremos $[T]_{\beta\beta'}$.

ii) Si T es un operador lineal $T : V \longrightarrow V$ y $\beta = \beta'$, entonces a la matriz A la denotaremos por $[T]_\beta$.

Ejemplo 20.

Sean $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por $T(x, y) = (2x-y, x+2y)$, $\beta = \{(1,-1), (2,2)\}$ y $\beta' = \{(2,1), (1,3)\}$. Calculemos la matriz A tal que $[T(x)]_{\beta'} = A[x]_\beta$.

Solución

Observemos que

$$T(1,-1) = (3,-1) \quad \text{y} \quad T(2,2) = (2,6)$$

Determinemos el vector de coeficientes $[\,T(\alpha_j)\,]_{\beta'}$ para cada $\alpha_j \in \beta$. Estos coeficientes satisfacen

$$\begin{aligned} T(1,-1) = (3,-1) &= l_1(2,1) + l_2(1,3) \\ T(2,2) = (2,6) &= l_3(2,1) + l_4(1,3) \end{aligned}$$

es decir,

$$\begin{aligned} 3 &= 2l_1 + l_2 \\ -1 &= l_1 + 3l_2 \end{aligned} \qquad \text{y} \qquad \begin{aligned} 2 &= 2l_3 + l_4 \\ 6 &= l_3 + 3l_4 \end{aligned}$$

Resolviendo este sistema de ecuaciones lineales obtenemos que $l_1 = 2,\ l_2 = -1;$ $l_3 = 0,\ l_4 = 2$. Por tanto, la matriz A es

$$A = \begin{bmatrix} 2 & 0 \\ -1 & 2 \end{bmatrix}$$

Ejemplo 21.
Sean $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ definida por $T(x,y) = (3x-y, 2x-2y, x+4y)$, $\beta = \{(1,-1),(0,2)\}$ y $\beta' = \{(1,1,0),(0,1,2),(-1,1,3)\}$. Calculemos la matriz A tal que $[\,T(x)\,]_{\beta'} = A[\,x\,]_\beta$.

Solución

A partir de la definición de $T(\cdot\,,\cdot)$ se tiene que

$$T(1,-1) = (4,4,-3) \quad \text{y} \quad T(0,2) = (-2,-4,8)$$

Los vectores de coeficientes $[\,T(\alpha_j)\,]_{\beta'}$ para cada $\alpha_j \in \beta$ satisfacen

$$\begin{aligned} T(1,-1) = (4,4,-3) &= l_1(1,1,0) + l_2(0,1,2) + l_3(-1,1,3) \\ T(0,2) = (-2,-4,8) &= l_4(1,1,0) + l_5(0,1,2) + l_6(-1,1,3) \end{aligned}$$

es decir,

$$\begin{aligned} 4 &= l_1 - l_3 \\ 4 &= l_1 + l_2 + l_3 \\ -3 &= 2l_2 + 3l_3 \end{aligned} \qquad \text{y} \qquad \begin{aligned} -2 &= l_4 - l_6 \\ -4 &= l_4 + l_5 + l_6 \\ 8 &= 2l_5 + 3l_6 \end{aligned}$$

Resolviendo este sistema de ecuaciones lineales, obtenemos que $l_1 = 7,\ l_2 = -6,\ l_3 = 3;\ l_4 = -14,\ l_5 = 22,\ l_6 = -12$. Por tanto, la matriz A es

$$A = \begin{bmatrix} 7 & -14 \\ -6 & 22 \\ 3 & -12 \end{bmatrix}$$

▲

Ahora queremos averiguar qué sucede con la matriz que representa una transformación lineal *si las bases cambian*. Aquí estudiaremos este problema sólo para operadores lineales $T : V \longrightarrow V$ $(\dim V = n)$ [2], y nuestro problema aparecerá entonces así: si β y β' son dos bases diferentes de V, ¿cuál es la relación entre $[\,T\,]_\beta$ y $[\,T\,]_{\beta'}$? Tenemos la respuesta en el siguiente teorema:

Teorema 6. ***(Matrices que representan a la misma transformación lineal)***
Si $T : V \longrightarrow V$ es un operador lineal y $\beta = \{\,\alpha_1, \ldots, \alpha_n\,\}$, y $\beta' = \{\,\alpha'_1, \ldots, \alpha'_n\,\}$ dos bases para V, entonces existe una matriz $P \in \mathfrak{M}_{n\times n}$ invertible tal que

$$[\,T\,]_{\beta'} = P^{-1}[\,T\,]_\beta P$$

Más aún, las columnas P_j de la matriz P están dadas por

$$P_j = [\,\alpha'_j\,]_\beta$$

Demostración
Sea $P = [\,P_1 \,|\, P_2 \,|\, \ldots \,|\, P_n\,]$, donde la columna P_j está definida por $P_j = [\,\alpha'_j\,]_\beta$. Entonces, para todo $\alpha \in V$ (ejercicio 3 de la sección 2 anterior) se tiene que,

$$[\,\alpha\,]_\beta = P[\,\alpha\,]_{\beta'} \tag{1}$$

y, por tanto,

$$[\,T(\,\alpha\,)\,]_\beta = P[\,T(\,\alpha\,)\,]_{\beta'} \tag{2}$$

Además, por definición,

$$[\,T(\,\alpha\,)\,]_\beta = [\,T\,]_\beta[\,\alpha\,]_\beta \tag{3}$$

Combinando (1), (2) y (3) se obtiene que

$$P[\,T(\,\alpha\,)\,]_{\beta'} = [\,T(\,\alpha\,)\,]_\beta = [\,T\,]_\beta[\,\alpha\,]_\beta = [\,T\,]_\beta P[\,\alpha\,]_{\beta'}$$

Luego

$$[\,T(\,\alpha\,)\,]_{\beta'} = [\,P^{-1}[\,T\,]_\beta P\,][\,\alpha\,]_{\beta'}$$

[2] Para el caso general de transformaciones lineales $T : V \to W$, ver el ejercicio 11 de esta sección.

Y así, podemos escribir que

$$[\,T\,]_{\beta'} = P^{-1}[\,T\,]_{\beta}P$$

con lo que se demuestra el teorema. ■

Ejemplo 22.
Sean $\beta = \{\,(\,1,2\,),(3,-1)\,\}$, $\beta' = \{\,(\,1,0\,),(0,1)\,\}$ y $T:\mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por $T(\,1,2\,) = 2(\,1,2\,)$, $T(\,3,-1\,) = 3(\,3,-1\,)$. Entonces, puesto que

$$\begin{aligned}(\,1,0\,) &= \frac{1}{7}(\,1,2\,) + \frac{2}{7}(\,3,-1\,)\\ (\,0,1\,) &= \frac{3}{7}(\,1,2\,) - \frac{1}{7}(\,3,-1\,)\end{aligned}$$

y la matriz de T relativa a la base β es

$$[\,T\,]_{\beta} = \begin{bmatrix} 2 & 0\\ 0 & 3\end{bmatrix}$$

entonces la matriz de T relativa a la base β' es

$$[\,T\,]_{\beta'} = P^{-1}[\,T\,]_{\beta}P = \begin{bmatrix} 1 & 3\\ 2 & -1\end{bmatrix}\begin{bmatrix} 2 & 0\\ 0 & 3\end{bmatrix}\begin{bmatrix} \frac{1}{7} & \frac{3}{7}\\ \frac{2}{7} & -\frac{1}{7}\end{bmatrix} = \begin{bmatrix} \frac{20}{7} & -\frac{3}{7}\\ -\frac{2}{7} & \frac{15}{7}\end{bmatrix}$$

Observemos que, efectivamente,

$$\begin{aligned}T(\,1,0\,) &= \frac{20}{7}(\,1,0\,) - \frac{2}{7}(\,0,1\,)\\ T(\,0,1\,) &= -\frac{3}{7}(\,1,0\,) + \frac{15}{7}(\,0,1\,)\end{aligned}$$

confirmando nuestro cálculo previo.

Definición 4. (Matrices similares (Cauchy (1826)))
Dos matrices A, $B \in \mathfrak{M}_{n\times n}$ se dicen *similares* si, y sólo si, existe una matriz invertible $P \in \mathfrak{M}_{n\times n}$ tal que

$$B = P^{-1}AP$$

Por lo tanto, de acuerdo con el teorema 6, *todas las matrices que representan (en alguna base) a un operador lineal, son similares.*

Ejemplo 23.

Si $A = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}$ y $B = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}$ son matrices similares y $P = c\begin{bmatrix} 4 & 1 \\ 1 & -1 \end{bmatrix}$, determinemos el valor de la constante c de tal manera que $P^{-1}AP = B$.

Solución

Observemos que si $c \neq 0$,

$$P^{-1}AP = \begin{bmatrix} \frac{1}{5c} & \frac{1}{5c} \\ \frac{1}{5c} & -\frac{4}{5c} \end{bmatrix}\begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 4c & c \\ c & -c \end{bmatrix} = \begin{bmatrix} \frac{1}{5c} & \frac{1}{5c} \\ \frac{1}{5c} & -\frac{4}{5c} \end{bmatrix}\begin{bmatrix} 24c & c \\ 6c & -c \end{bmatrix} = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}$$

así, *cualquier* matriz de la forma $\begin{bmatrix} 4c & c \\ c & -c \end{bmatrix}$ (con $c \neq 0$) sirve para establecer la similaridad de las matrices A y B. ¿Qué significa esto geométricamente?

Ejemplo 24.

a) $A = \begin{bmatrix} 7.3 & 0.2 & -3.7 \\ -11.5 & 1.0 & 5.5 \\ 17.7 & 1.8 & -9.3 \end{bmatrix}$ es similar a $P^{-1}AP = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ para

$P = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}$ y $P^{-1} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix}$ como

fácilmente lo puede comprobar el lector.

b) Puesto que $\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$,

entonces $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ y $\begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$ son similares. Aquí debemos notar

que $\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$.

a. El rango de una matriz

En 1878, Ferdinand George Frobenius [1849-1917] escribió un importante trabajo sobre matrices llamado *On Linear Substitutions and Bilinear Forms*. En este artículo incluye la importante definición de *rango de una matriz*, que aquí presentamos.

Definición 5. (Rango columna de una matriz (Frobenius (1878)))
Si $A \in \mathfrak{M}_{m\times n}$, definimos su *rango (columna)*, que denotamos por $\rho(A)$, como

$$\rho(A) = \dim\{ A_1, A_2, \cdots , A_n \}$$

donde escribimos A de la siguiente manera:

$$A = [\ A_1 \,|\, A_2 \,|\, \cdots \,|\, A_n\ ]$$

Es decir, *el rango de una matriz es el máximo número de sus columnas linealmente independientes.*

Teorema 7. ***(Propiedades del rango de una matriz)***
Si A es una matriz $m \times n$, entonces se tiene lo siguiente:

a) $\rho(A) = \rho(A^T)$ *y* $\rho(AA^T) = \rho(A^T A) = \rho(A)$

b) *Si $r(A)$ es el número de filas linealmente independientes (que también se le conoce como rango fila de A), entonces*

$$r(A) = \rho(A)$$

c) *Si $m = n$, entonces A es invertible si, y sólo si, $\rho(A) = n$.*

d) $\rho(AB) \leq \min\{ \rho(A), \rho(B) \}$.

e) $\rho(A) \leq \min\{ m, n \}$.

f) *Si B es invertible, entonces $\rho(AB) = \rho(A)$; y si C es invertible, entonces $\rho(CA) = \rho(A)$.*

g) *La operaciones fila sobre una matriz no alteran su rango (y, por tanto, el método gaussiano es un algoritmo para encontrar el rango de una matriz).*

h) $\rho(A + B) \leq \rho(A) + \rho(B)$.

i) *Las matrices similares tienen el mismo rango.*

Demostración

a) i) Veamos que $\rho(A) = \rho(A^T)$ sólo en el caso 2×2. El caso puede extenderse fácilmente a matrices $m\times n$. Podemos asumir que $\rho(A) = 1$ pues en otro caso ($\rho(A) = 2$) es inmediato. Por tanto, para algún $k \in \mathbb{R}$, se tiene que

$$A = \begin{bmatrix} a_{11} & ka_{11} \\ a_{21} & ka_{21} \end{bmatrix}$$

donde asumiremos (sin pérdida de generalidad) que $a_{11} \neq 0$ (es decir, si las columnas de A son linealmente dependientes).

Entonces, también, para algún $\lambda \in \mathbb{R}$ se tiene que

$$A = \begin{bmatrix} a_{11} & a_{12} \\ \lambda a_{11} & \lambda a_{12} \end{bmatrix}$$

pues para esto sólo basta tomar $\lambda = \dfrac{a_{21}}{a_{11}}$, pues en tal caso, $a_{21} = \lambda a_{11}$ y $a_{22} = ka_{21} = \dfrac{a_{12}}{a_{11}} a_{21} = \dfrac{a_{21}}{a_{11}} a_{12} = \lambda a_{12}$. Así, $\rho(A^T) = 1$.

ii) Que $\rho(AA^T) = \rho(A) = \rho(A^T A)$ se ve del hecho de que las columnas de la matriz AA^T son combinaciones lineales de las columnas de la matriz A, y viceversa. De manera similar, para la matriz $A^T A$ sus columnas son combinaciones lineales de la matriz A^T y viceversa; el resultado se tiene entonces del hecho de que $\rho(A) = \rho(A^T)$.

b) Es consecuencia directa del literal a) pues las filas de A son columnas de A^T.

c) Sabemos que A es invertible si, y sólo si, las n columnas de A son linealmente independientes; es decir, si, y sólo si, $\rho(A) = n$.

d) Si A es una matriz $m \times n$ y B es una matriz $n \times p$, entonces cada columna de AB será una combinación lineal de las n columnas de A y, por tanto, el rango de AB no puede ser mayor que el rango de A. De forma similar, podemos decir que el rango de $(AB)^T = B^T A^T$ no puede ser mayor que el rango de B^T. Pero por el literal a), el rango de $(AB)^T$ es el mismo rango de AB, y el rango de B^T es el mismo rango de B.

e) En el literal d), tómese $B = I_n$. Entonces $\rho(B) = n$.

f) Por el literal d), $\rho(AB) \leq \rho(A)$ y $\rho(CA) \leq \rho(A)$; pero $A = (AB)B^{-1}$ y $A = C^{-1}(CA)$; luego, por el mismo literal d), $\rho(A) \leq \rho(AB)$ y $\rho(A) \leq \rho(CA)$; así, $\rho(AB) = \rho(A) = \rho(CA)$.

g) En efecto, como el rango de una matriz es, por el literal b) anterior, igual al máximo número de filas linealmente independientes y, por tanto, igual a la dimensión del espacio generado por ellas, las posibles combinaciones lineales de estas filas no alteran tal dimensión.

h) Si $\beta_1 = \{A_1, A_2, ..., A_{\rho(A)}\}$ es una base para el espacio columna de A, y $\beta_2 = \{B_1, B_2, ..., B_{\rho(B)}\}$ es una base para el espacio columna de B, entonces en

$$\beta = \beta_1 \cup \beta_2$$

existe una base para el espacio columna de $A+B$, que puede no ser todo el conjunto $\beta_1 \cup \beta_2$; es decir,

$$\rho(A+B) \leq \rho(A) + \rho(B)$$

i) Sean A y B dos matrices similares; es decir, existe una matriz invertible P tal que $B = P^{-1}AP$. Entonces el literal f) de este teorema implica que $\rho(\,B\,) = \rho(\,P^{-1}AP\,) = \rho(\,AP\,) = \rho(\,A\,)$. ■

Ahora podemos enunciar un resultado que podría esperarse: el rango de una transformación lineal *coincide* con el rango de *cualquiera* de las matrices que la representan.

Teorema 8.
Si $T : V \longrightarrow V$ (con $\dim V = n$) es una transformación lineal y A es una matriz asociada cualquiera, entonces

$$\rho(\,T\,) = \rho(\,A\,)\,;$$

es decir, el rango de la transformación lineal y de la matriz coinciden.

Demostración

a) En primer lugar, si $A_{n\times n}$ representa a la transformación T, entonces $\mathrm{Im}(\,T\,)$ es el espacio generado por las columnas de A y, por tanto, $\rho(\,T\,) = \rho(\,A\,)$.

b) En segundo lugar, si $B = P^{-1}AP$ para cierta matriz invertible P, entonces $\rho(\,B\,) = \rho(\,A\,)$ por el literal i) del teorema 7. ■

Definición 6. (Rango completo)
Si A es una matriz $m \times n$, diremos que A tiene *rango completo* (*full rank*) si, y sólo si,

$$\rho(\,A\,) = \text{mín}\{\,m, n\,\}$$

Claramente, una *matriz cuadrada* tiene rango completo si, y sólo si, es invertible.

Ejemplo 25.
El rango de la matriz

$$A = \begin{bmatrix} 0 & 1 & 5 & 0 \\ 0 & 2 & 10 & 0 \end{bmatrix}$$

es $\rho(\,A\,)$=1 pues $\rho(\,A\,) = \rho(\,A^T\,)$ y las filas de A satisfacen $(\,0,1,5,0\,) = \frac{1}{2}(\,0,2,10,0\,)$. Por lo tanto, esta matriz no tiene rango completo. ¿Cuántas soluciones tendrá el sistema lineal $Ax = 0$ para $x \in \mathbb{R}^4$?

Ejemplo 26.
Encontremos el rango de la matriz 4×4

$$A = \begin{bmatrix} 1 & 2 & -3 & 4 \\ -2 & 1 & 7 & -5 \\ 2 & 5 & -3 & 4 \\ 2 & 10 & -2 & 7 \end{bmatrix}$$

¿Cuántas soluciones tendrá el sistema lineal $Ax = 0$ para $x \in \mathbb{R}^4$? ¿Tiene A rango completo?

Solución

Realizando operaciones elementales de filas sobre la matriz A se tiene que

$$\begin{bmatrix} 1 & 2 & -3 & 4 \\ -2 & 1 & 7 & -5 \\ 2 & 5 & -3 & 4 \\ 2 & 10 & -2 & 7 \end{bmatrix} \begin{matrix} F_1 \\ F_2 \\ F_3 \\ F_4 \end{matrix}$$

$$\begin{bmatrix} 1 & 2 & -3 & 4 \\ 0 & 5 & 1 & 3 \\ 0 & 1 & 3 & -4 \\ 0 & 6 & 4 & -1 \end{bmatrix} \begin{matrix} \\ F_2 \longleftrightarrow F_2 + 2F_1 \\ F_3 \longleftrightarrow F_3 - 2F_1 \\ F_4 \longleftrightarrow F_4 - 2F_1, \end{matrix}$$

$$\begin{bmatrix} 1 & 2 & -3 & 4 \\ 0 & 1 & -11 & 19 \\ 0 & 1 & 3 & -4 \\ 0 & 6 & 4 & -1 \end{bmatrix} \quad F_2 \longleftrightarrow F_2 - 4F_3$$

$$\begin{bmatrix} 1 & 0 & 19 & -34 \\ 0 & 1 & -11 & 19 \\ 0 & 0 & 14 & -23 \\ 0 & 0 & 70 & -115 \end{bmatrix} \begin{matrix} F_1 \longleftrightarrow F_1 - 2F_2 \\ \\ F_3 \longleftrightarrow F_3 - F_2 \\ F_4 \longleftrightarrow F_4 - 6F_2 \end{matrix}$$

$$\begin{bmatrix} 1 & 0 & 19 & -34 \\ 0 & 1 & -11 & -19 \\ 0 & 0 & 1 & -\frac{23}{14} \\ 0 & 0 & 70 & -115 \end{bmatrix} \quad F_3 \longleftrightarrow \frac{1}{14}F_3$$

$$\begin{bmatrix} 1 & 0 & 0 & -\frac{39}{14} \\ 0 & 1 & 0 & \frac{13}{14} \\ 0 & 0 & 1 & -\frac{23}{14} \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} F_1 \longleftrightarrow F_1 - 19F_3 \\ F_2 \longleftrightarrow F_2 + 11F_3 \\ \\ F_4 \longleftrightarrow F_4 - 70F_3 \end{matrix}$$

Por tanto, $\rho(A) = 3$ y el sistema $Ax = 0$ con $x \in \mathbb{R}^4$ tendrá infinitas soluciones. Observemos que $\rho(A) = 3 \neq 4 = \text{mín}\{m, n\}$ y, por tanto, A *no* tiene rango completo.

Ejemplo 27.
Encontremos el rango de

$$A = \begin{bmatrix} 4 & 3 & -1 \\ 3 & -2 & 12 \end{bmatrix}$$

¿Tiene A rango completo?

Solución
Aquí $\rho(A) = 2$, pues las *filas* de A, $(4, 3, -1)$ y $(3, -2, 12)$, son linealmente independientes. Observemos que éste coincide con el rango columna de A pues

$$\begin{pmatrix} -1 \\ 12 \end{pmatrix} = 2\begin{pmatrix} 4 \\ 3 \end{pmatrix} - 3\begin{pmatrix} 3 \\ -2 \end{pmatrix}$$

y $\begin{pmatrix} 4 \\ 3 \end{pmatrix}$ y $\begin{pmatrix} 3 \\ -2 \end{pmatrix}$ son linealmente independientes. Claramente, $\text{mín}\{m, n\} = 2 = \rho(A)$ y, por tanto, A tiene rango completo. ¿Cuántas soluciones tiene entonces el sistema $Ax = 0$?

Ejemplo 28.
Encontremos el rango de

$$A = \begin{bmatrix} 1 & 5 & 7 \\ 2 & 3 & 1 \end{bmatrix}$$

¿Tiene A rango completo?

Solución
Aquí $\rho(A) = 2$ pues las filas de A, $(1, 5, 7)$ y $(2, 3, 1)$, son linealmente independientes Así, $\rho(A)$=min $\{2, 3\}$, y por tanto A tiene rango completo.

Teorema 9. ***(¿Qué hemos aprendido sobre los sistemas homogéneos?)***
Sea A una matriz $n \times n$; entonces las seis afirmaciones siguientes son equivalentes:

a) $AX = 0$ *tiene solución única* $X = 0$.

b) A *es invertible.*

c) $\det A \neq 0$.

d) *Las columnas de A forman una base para $\mathbb{R}^n$. También las filas de A forman una base para $\mathbb{R}^n$.*

e) *La dimensión del espacio de soluciones del sistema homogéneo $AX = 0$ es igual a cero.*

f) $\rho(A) = n$.

Ejercicios 3

1) Encuentre una matriz asociada a los siguientes operadores lineales:

a) $T(x, y) = (x + y, 2x - 3y)$

b) $T(x, y, z) = (x + z, y - z, 3x - y - 2z)$

c) $$\begin{aligned} y_1 &= 2x_1 + 4x_2 + x_3 \\ y_2 &= x_1 + x_2 - x_3 \\ y_3 &= 4x_1 - 4x_2 + 3x_3 \end{aligned}$$

d) $$\begin{aligned} y_1 &= \frac{1}{2}x_1 - \frac{1}{3}x_3 \\ y_2 &= x_1 + x_2 + x_3 \\ y_3 &= \frac{1}{4}x_1 + \frac{1}{7}x_3 \end{aligned}$$

2) Si T es el operador lineal de $\mathbb{R}^2$ en $\mathbb{R}^3$ definido por

$$T(x, y) = (x + y, x - 2y, 4x + 3y)$$

a) ¿Cuál es la matriz de T en la base canónica?

b) ¿Cuál es la matriz de T en la base $\{(3, 2), (1, 3)\}$?

3) Si T es el operador sobre $\mathbb{R}^3$ definido por

$$T(x, y, z) = (2x + y, x - y - z, 4x + 2y + z)$$

a) ¿Cuál es la matriz de T en la base canónica?

b) ¿Cuál es la matriz de T en la base $\{(1, 0, 1), (1, 2, 3), (-1, 4, 0)\}$?

c) ¿Cuál es la matriz P que muestra la similaridad entre estas dos matrices?

d) Interprete geométricamente lo anterior.

4) Muestre que los siguientes pares de matrices son similares:

a) $A = \begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}$

b) $A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 8 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

Interprete aquí, geométricamente, la condición de que estas matrices sean similares.

5) Sea A una matriz $m \times n$ donde $n \leq m$. Pruebe que A tiene rango completo si, y sólo si, $A^T A$ es invertible.

6) Encuentre el rango de las siguientes matrices calculando primero a través de sus columnas y después a través de sus filas (y corroborando que coincidan):

a) $A = \begin{bmatrix} 0 & 1 & 3 & 2 \\ 1 & 0 & 1 & 5 \\ 2 & 2 & 0 & 3 \end{bmatrix}$ b) $A = \begin{bmatrix} 3 & 1 \\ 4 & 1 \\ 0 & 1 \end{bmatrix}$

c) $A = \begin{bmatrix} 1 & 2 & 2 \\ 3 & 5 & 4 \\ -2 & 7 & 2 \end{bmatrix}$ d) $A = \begin{bmatrix} 2 & 1 & 0 & 1 & 2 \\ 1 & 1 & 0 & 1 & -1 \\ 7 & 2 & 3 & -2 & 1 \\ 0 & -1 & 0 & -1 & 0 \end{bmatrix}$

¿Cuáles de ellas tienen rango completo?

7) En cada caso de los cuatro anteriores, verifique que si $T(x) = Ax$, entonces $\mathrm{nul}(T) + \rho(T) = n$. ¿Cuál es una base para $\mathrm{Nu}(T)$?

8) Muestre que el rango de la matriz

$$A = \begin{bmatrix} 3 & 0 & 2 & 2 \\ -6 & 42 & 24 & 54 \\ 21 & -21 & 0 & -15 \end{bmatrix}$$

es igual a 2.

9) Resolviendo el sistema

$$\begin{aligned} 2x - 3y + z &= 0 \\ x + y - z &= 0 \end{aligned}$$

a) Encuentre el núcleo de la transformación lineal $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definida por

$$T(x, y, z) = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \left(\equiv A \begin{bmatrix} x \\ y \\ z \end{bmatrix} \right)$$

b) ¿Cuál es la dimensión de la imagen de T?

c) ¿Cuál es el rango de A?

d) ¿Coinciden ambos números? Explique.

10) Resolviendo el sistema

$$\begin{aligned} 2x - 3y + z &= 0 \\ x + y - z &= 0 \\ 3x + 4y &= 0 \\ 5x + y + z &= 0 \end{aligned}$$

a) Encuentre el núcleo de la transformación lineal $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ definida por

$$T(\,x, y, z\,) = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 1 & -1 \\ 3 & 4 & 0 \\ 5 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \left(\equiv A \begin{bmatrix} x \\ y \\ z \end{bmatrix} \right)$$

b) ¿Cuál es la dimensión de la imagen de T?

c) ¿Cuál es el rango de A?

d) ¿Coinciden ambos números?

**11) Pruebe que si $T : V \rightarrow W$ es una transformación lineal con $dimV = n$ y $dimW = m$, y además β_1 y β_2 son bases de V, y γ_1 y γ_2 bases de W, entonces existen dos matrices inversibles, P de tamaño $n \times n$ y Q de tamaño $m \times m$, tales que $[T]_{\beta_2\gamma_2} = Q^{-1}[T]_{\beta_1\gamma_1}P$ donde Q es la matriz de paso de γ_1 a γ_2, y P es la matriz de paso de β_1 a β_2.

4. Estructura de los conjuntos de transformaciones lineales

Un hecho profundamente fundamental en la teoría del álgebra lineal es que al estudiar las transformaciones lineales de un espacio vectorial V en otro W, resulta que éstas "heredan" cierta estructura algebraica que encontraremos muy familiar, pues ya la habíamos visto en la lección 2 en el escenario de las *matrices*. Por tal razón, adelantaremos esta sección de una manera un tanto esquemática.

Teorema 10. ***(Suma y producto por escalar de transformaciones lineales (Peano (1888)))***
Sea $L(\,V, W\,) = \{\,T : V \longrightarrow W \,/\, T \text{ es una transformación lineal}\}$. Entonces $L(\,V, W\,)$ es un espacio vectorial donde las operaciones están definidas así:

a) *Si* $T,\ U \in L(V, W)$, *entonces, para* $x \in V$,

$$(T + U)(x) = T(x) + U(x)$$

b) *Si* $k \in \mathbb{R}$, *entonces, para* $x \in V$,

$$(kT)(x) = k\,T(x)$$

Demostración

a) Probemos primero que si $T,\ U \in L(V, W)$, entonces $(T+U) \in L(V, W)$. En efecto, para $x,\ y \in V,\ k \in \mathbb{R}$

i)
$$\begin{aligned}(T + U)(x + y) &= T(x + y) + U(x + y)\\ &= [T(x) + T(y)] + [U(x) + U(y)]\\ &= [T(x) + U(x)] + [T(y) + U(y)]\\ &= (T + U)(x) + (T + U)(y)\end{aligned}$$

ii) $(T + U)(kx) = T(kx) + U(kx) = kT(x) + kU(x) = k[T(x) + U(x)]$

b) En segundo lugar, probemos que si $T \in L(V, W)$, entonces, para $k \in \mathbb{R}$, $kT \in L(V, W)$; en efecto, para $x,\ y \in V,\ l \in \mathbb{R}$

i) $(kT)(x + y) = kT(x + y) = k[T(x) + T(y)] = kT(x) + kT(y)$

ii) $(kT)(lx) = kT(lx) = klT(x) = l(kT)(x)$ ∎

Teorema 11. ***(Propiedades de la suma y producto por escalar de transformaciones lineales)***
Si T, U *y* $Z \in L(V, W)$, *entonces, para todo* $x \in V$,

a) $(T + U)(x) = (U + T)(x)$ *(ley conmutativa).*

b) $((T + U) + Z)(x) = (T + (U + Z))(x)$ *(ley asociativa).*

c) *Existe una transformación* $O \in L(V, W)$, *que llamaremos transformación cero, tal que* $O(x) = 0$ *para todo* $x \in V$ *y que satisface que* $T + O = O + T = T$ *para toda* $T \in L(V, W)$.

d) *Para cada* $T \in L(V, W)$, *existe la transformación* $-T \in L(V, W)$ *tal que* $T(x) + (-T(x)) = 0$ *para todo* $x \in V$.

Además, para todo par de escalares $k,\ l \in R$ *se tiene que*

e) $k\,(l\,T(x)) = kl\,T(x)$

f) $k\,(\,T(\,x\,) + U(\,x\,)\,) = k\,T(\,x\,) + k\,U(\,x\,)$

g) $(\,k + l\,)T(\,x\,) = k\,T(\,x\,) + l\,T(\,x\,)$

Demostración

Se deja como ejercicio (sencillo) para el lector. ■

Teorema 12.
Si $\dim V = n$ *y* $\dim W = m$*, entonces* $L(\,V, W\,)$ *tiene dimensión* mn*.*

Demostración

Tomemos $\beta = \{\,\alpha_1,\,\ldots,\,\alpha_n\,\}$ y $\beta' = \{\,\beta_1,\,\ldots,\,\beta_m\,\}$ como bases ordenadas para V y W respectivamente. Sean, además, $\{\,T_{p,q}\,\}_{\substack{1\leq p\leq m\\ 1\leq q\leq n}}$ la familia de transformaciones de V en W definidas por

$$T_{p,q}(\,\alpha_i\,) = \begin{cases} 0 & \text{si} \quad i \neq q \\ \beta_p & \text{si} \quad i = q \end{cases}$$

Veamos que estas mn funciones generan $L(\,V, W\,)$ y que, además, son linealmente independientes.

i) $\{T_{p,q}(\cdot)\}$ genera a $L(V, W)$: Sea $T : V \longrightarrow W$ una transformación cualquiera. Escribiendo
$$T = \sum_{\substack{1\leq p\leq m\\ 1\leq q\leq n}} k_{pq} T_{pq}$$
para ciertos $k_{pq} \in \mathbb{R}$, y evaluando en α_i para $i = 1, 2, ..., n$, encontramos que
$$T(\alpha_i) = \sum_{\substack{1\leq p\leq m\\ 1\leq q\leq n}} k_{pq} T_{pq}(\alpha_i) = \sum_{1\leq p\leq m} k_{pi}\beta_p \ ;$$
de manera que los $\{k_{pq}\}$ estarán definidos como los coeficientes que acompañan a los vectores $\{\beta_p\}_{1\leq p\leq m}$ en la combinación lineal de los $\{T(\alpha_q)\}_{0\leq q\leq n}$.

ii) $\{T_{p,q}(\cdot)\}$ es linealmente independiente pues si

$$\sum_{\substack{1\leq p\leq m\\ 1\leq q\leq n}} k_{pq} T_{pq} = 0$$

entonces, evaluando nuevamente en α_i para $1 \leq i \leq n$, se tendría que

$$\sum_{\substack{1\leq p\leq m\\1\leq q\leq n}} k_{pq}T_{pq}(\alpha_i)=0$$

y, por tanto, $\sum_{1\leq p\leq m} k_{pi}\beta_p = 0$, lo que implica que $k_{pi} = 0$ para todo $1 \leq p \leq m$. Como esto es cierto para todo $1 \leq i \leq n$, la independencia lineal está mostrada. ■

Pero el espacio vectorial $L(V,W)$ tiene una estructura algebraica que va más allá de las *características lineales* de un espacio vectorial. De hecho, si existe el producto entre matrices, es de esperarse que también exista el "producto" de transformaciones lineales, que no es más que la *composición* de transformaciones lineales a la manera de la composición de funciones ordinarias (Volumen 0 (Fundamentos)).

Teorema 13. ***(Composición de transformaciones lineales)***
Si tenemos $T \in L(V,W)$ y $U \in L(W,Z)$, entonces la composición (o producto) de transformaciones $UT : V \longrightarrow Z$ definida como es usual por $((UT)(x) = U(T(x)))$ es también una transformación lineal.

Demostración

a) *Si $x,\ y \in V$, entonces* $UT(x+y) = U(T(x+y)) =$

$$U(T(x)+T(y)) = U(T(x)) + U(T(y)) = UT(x) + UT(y)$$

b) *Si $k \in \mathbb{R},\ x \in V$*

$$(UT)(kx) = U(T(kx)) = U(kT(x)) = kU(T(x)) = k(UT(x)) \quad ■$$

Teorema 14. ***(Propiedades del producto de transformaciones lineales)***
Si T, U y $Z \in L(V,W)$ entonces para todo $x \in V$,

a) $UT(x) \neq TU(x)$ (no se satisface la ley conmutativa) [3]

[3] Conociendo la estrecha conexión entre transformaciones lineales y matrices, y aceptando, en principio, la intuición geométrica de que toda transformación lineal es una combinación de "movimientos rígidos", ahora podemos entender que la multiplicación de dos matrices no es más que la "composición de dos movimientos rígidos", y estos, no conmutan. Por esta razón, la multiplicación de dos matrices no puede ser conmutativa. Si al lector no le es claro, en este punto, que dos "movimiento rígidos" no conmutan, le invitamos a realizar el siguiente ejercicio lúdico: en su computador dibuje una elipse estándar en el plano cartesiano (por ejemplo, tome $\frac{x^2}{3^2}+\frac{y^2}{4^2}=1$) y

b) $U(TZ(x)) = UT(Z(x))$ *(ley asociativa)*

c) $U((T(x) + Z(x)) = U(T(x)) + U(Z(x))$ *(ley distributiva a derecha)*

d) $(T + Z)(U(x)) = T(U(x)) + Z(U(x))$ *(ley distributiva a izquierda)*

e) $k(U(T(x))) = (kU(T(x))) = (U(kT(x)))$, *donde* $k \in \mathbb{R}$

Demostración

Todas son consecuencias inmediatas de la linealidad de las transformaciones T, U, Z. ■

Definición 7. (Transformación invertible)
Decimos que $T : V \longrightarrow V$ es *invertible* si existe una transformación $U : V \longrightarrow V$ tal que $TU = UT = I$ (transformación lineal idéntica de V en V.).

A esta U se le conoce como la *transformación inversa* de T y se le notará por T^{-1}. ¿Podría el lector justificar esta notación mostrando que esta transformación es única?

Teorema 15. ***(Propiedades de la transformación inversa)***
Si T *y* $U \in L(V, V)$ *son inversibles, entonces*

a) *La transformación inversa de* T *es única.*

b) T^{-1} *también es invertible y además* $(T^{-1})^{-1} = T$.

c) TU *es invertible y además* $(TU)^{-1} = U^{-1}T^{-1}$.

d) $T + U$ *no es necesariamente invertible.*

e) *Si* $TC = TD$ *para ciertas transformaciones* C *y* D, *entonces* $C = D$.

Demostración

Es directa, y se deja como ejercicio para el lector. ■

---

realice las siguientes operaciones:

a) En primer lugar, rote la elipse 45°, y después cambie la escala del eje Y al doble de la escala original. Observe la figura que aparece.

b) Tome de nuevo la elipse original, y cambie primero la escala del eje Y al doble de la escala original; luego rote la figura 45°. Observe la figura que aparece.

c) Corrobore que las dos figuras en a) y b) son diferentes.

Ejercicios 4

1) En este punto entendemos que las matrices están íntimamente conectadas con los sistemas de ecuaciones lineales (lecciones 2 y 3), y también entendemos que las matrices son, fundamentalmente, otra forma de describir las transformaciones lineales (lección 6 presente). Entonces ¿cuál será la conexión *geométrica* entre los sistemas de ecuaciones lineales y las transformaciones lineales?

2) Si $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por $T(\,x\,) = Ax$, donde

$$A = \begin{bmatrix} \cos\,\theta & -\,\text{sen}\,\theta \\ \text{sen}\,\theta & \cos\,\theta \end{bmatrix}$$

es la transformación que representa una rotación de ejes en un ángulo θ, ¿qué significa, geométricamente, la igualdad

$$A^n = \begin{bmatrix} \cos\,(\,n\theta\,) & -\,\text{sen}\,(\,n\theta\,) \\ \text{sen}\,(\,n\theta\,) & \cos\,(\,n\theta\,) \end{bmatrix}$$

para cualquier $n \in \mathbb{N}$?

3) Demuestre los teoremas 11, 14 y 15.

5. Isomorfismos

La idea fundamental de los isomorfismos entre espacios es encontrar qué espacios *preservan* la estructura lineal; o, en otra forma, qué espacios son "linealmente similares". En esta sección demostraremos que todos los espacios finito-dimensionales son "linealmente similares" a algún espacio cartesiano $\mathbb{R}^n$. Es decir, *los espacios* $\mathbb{R}^n$ *son los "únicos" espacios vectoriales finito-dimensionales.*

Definición 8. (Espacios vectoriales isomorfos)
Diremos que dos espacios vectoriales V y W son *isomorfos* si, y sólo si, existe una transformación lineal *biyectiva* $T : V \longrightarrow W$. En tal caso, a la transformación T se le llama un *isomorfismo entre* V *y* W.

Teorema 16.
Si $\dim V = n$ *y* $\dim W = m$*, entonces* $L(\,V, W\,)$ *es isomorfo a* $\mathfrak{M}_{m\times n}$ *(y ambos isomorfos a* $\mathbb{R}^{mn}$*). En particular,* $L(\,\mathbb{R}^n, \mathbb{R}^m\,)$ *es isomorfo a* $\mathfrak{M}_{m\times n}$.

Demostración

Sea β una base fija de V, y β' una base fija de W. Entonces

$$\begin{aligned} T : L(V,W) &\longrightarrow \mathfrak{M}_{m\times n} \\ U &\longrightarrow T(U) \end{aligned}$$

es un isomorfismo, donde $T(U)$ es la única matriz de la transformación lineal T relativa a las bases β y β' (teorema 5). ■

El siguiente teorema muestra, en definitiva, que:

a) Linealmente, las transformaciones y las matrices son "el mismo espacio".

b) Los "únicos" espacios vectoriales finito-dimensionales son los espacios $\mathbb{R}^n$.

Teorema 17. ***(Todos los espacios vectoriales de igual dimensión son isomorfos)***
Todo espacio vectorial V de dimensión n es isomorfo a $\mathbb{R}^n$.

Demostración

Sea $\beta = \{\alpha_1, \ldots, \alpha_n\}$ una base para V; definamos $T : V \longrightarrow \mathbb{R}^n$ tal que $T(\alpha_i) = e_i$, $i = 1, 2, \ldots, n$, donde $e_i = (0, \ldots, 1, \ldots, 0)$ es el i-ésimo vector canónico de $\mathbb{R}^n$.

a) Que T es lineal, es evidente.

b) Para probar que T es biyectiva, es suficiente probar que $\mathrm{Nu}(T) = \{0\}$. Para ello, sea $x \in V,\ x = \sum_{i=1}^{n} k_i\alpha_i,\ k_i \in \mathbb{R}$; entonces

$$T(x) = \sum_{i=1}^{n} k_i T(\alpha_i) = \sum_{i=1}^{n} k_i\, e_i = (0, \ldots, 0)$$

si, y sólo si, $k_i = 0$ para todo i; luego $x = 0$. Así que, efectivamente, $\mathrm{Nu}(T) = \{0\}$. De esta manera, T es biyectiva y, por tanto, un isomorfismo. ■

Ejemplo 29. (P_n es isomorfo a $\mathbb{R}^{n+1}$)

Podemos observar, de forma explícita, que P_n, el conjunto de todos los polinomios de grado menor o igual a n, es isomorfo a $\mathbb{R}^{n+1}$. Un isomorfismo casi evidente es el siguiente: si $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1} + a_nx^n$, entonces

$$\begin{aligned} T : P_n &\longrightarrow \mathbb{R}^{n+1} \\ p(x) &\longleftrightarrow (a_0, a_1, \ldots, a_{n-1}, a_n) \end{aligned}$$

satisface las condiciones de ser una transformación lineal y biyectiva.

Ejemplo 30. ($\mathbb{C}$ es isomorfo a $\mathbb{R}^2$)
Podemos construir un isomorfismo explícito entre el conjunto de los números complejos $\mathbb{C}$ (Volumen 0 (Fundamentos)) y el espacio cartesiano $\mathbb{R}^2$. En efecto, a cada número complejo $z = a+ib$ le asociamos el vector (a, b) en $\mathbb{R}^2$. Que esta función es un isomorfismo entre estos dos espacios vectoriales es un ejercicio simple para el lector.

Nota 3. (¿Y los espacios infinito-dimensionales?)
Ahora podría ser claro que un curso típico de álgebra lineal como el presentado en este texto sólo trata con espacios vectoriales finito-dimensionales, y que los "únicos" espacios vectoriales finito-dimensionales son los espacios euclidianos $\mathbb{R}^n$. Pero, ¿qué sucede con el estudio de los espacios infinito-dimensionales?

A comienzos del siglo XX, algunas de las investigaciones surgidas de las necesidades de la mecánica cuántica en física, condujeron a la creación de lo que hoy se conoce como *análisis funcional* que, en principio, es el estudio de ciertos espacios vectoriales infinito-dimensionales conocidos como *espacios de Hilbert* y, más generalmente, *espacios de Banach.* Los espacios de Hilbert (que son, fundamentalmente, espacios infinito-dimensionales donde hay definido un producto interior entre sus elementos a la manera del producto punto en $\mathbb{R}^n$ que satisface ciertas propiedades) fueron introducidos axiomáticamente por John von Neumann en 1929, aunque ya mucho antes se habían estudiado (y aplicado) ejemplos de ellos. Stephan Banach [1892-1945], por su parte, en su tesis doctoral de 1920 introdujo una teoría axiomática de los espacios vectoriales normados (es decir, de los espacios vectoriales infinito-dimensionales en los que se ha definido una *norma* para cada uno de los elementos, a la manera en que se definió la norma de vectores en $\mathbb{R}^n$). Hoy en día, el análisis funcional continúa su desarrollo teórico y aplicado, y es de fundamental interés desde el punto de vista del análisis matemático. Sobre los fundamentos del análisis funcional discutiremos en la lección 4 del Volumen 3 (Optimización y dinámica).

Ejercicios 5

1. ¿Podría el lector dar ejemplos de otros isomorfismos entre los espacios vectoriales finito-dimensionales discutidos previamente en la lección 5?

2. Probar que si P es una matriz $n \times n$ invertible, entonces

$$\begin{aligned} T : \mathfrak{M}_{n\times n} &\longrightarrow \mathfrak{M}_{n\times n} \\ A &\longrightarrow P^{-1}AP \end{aligned}$$

es un isomorfismo. ¿Cómo expresa, *geométricamente*, este resultado?

6. Contexto económico

a. El modelo de equilibrio general de Von Neumann (1932)

Tres años antes de que Wald publicara su prueba sobre la existencia de un equilibrio en una economía Walras-Cassel, el matemático húngaro John von Neumann presentó uno de los modelos lineales mejor logrados de la historia de la economía matemática. Su artículo *A Model of General Economic Equilibrium* de 1946 fue discutido por primera vez en 1932, y después publicado en alemán bajo el título *Über ein Ökonomisches Gleichungssystem und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes* en 1937.

En este artículo, Von Neumann va, incluso, más allá (en algunos aspectos importantes) que el modelo de Cassel. Y, notablemente, por primera vez en la literatura de la economía matematizada, aparece un teorema de existencia de equilibrios utilizando un teorema de punto fijo: *el teorema de punto fijo de Brouwer*, que más tarde utilizaría John Nash (1950) para garantizar la existencia de equilibrios en la teoría de juegos; y Kenneth Arrow y Gerard Debreu (1954)[4] para garantizar la existencia de un equilibrio walrasiano (aunque estos últimos también recurrieron a una versión más sofisticada del teorema de Brouwer, conocido como el *teorema de punto fijo de Kakutani* (1941)[5].

Von Neumann describe una economía en expansión caracterizada por una *producción lineal* en la cual *todos* los productos sirven de insumos a posteriores procesos productivos. La función de producción procesa insumos en productos, todos en *proporciones fijas*. Además, asume que cada mercancía es, o un insumo, o un producto de *todos* los procesos; el sector consumo se describe mediante un proceso donde bienes finalizados se utilizan como insumos en la producción de trabajo. *Así, el consumo es aquí un fenómeno tecnológico y, por tanto, las relaciones de demanda walrasianas desaparecen en este modelo.*

Cuando la economía se expande en períodos, von Neumann asume que no hay un límite a la oferta de tierra, mano de obra, u otros factores que pongan fin a la expansión. Y se pregunta si existe una *tasa constante de crecimiento* que dé beneficios nulos (como corresponde a la competencia perfecta: beneficios positivos atraen firmas al mercado haciendo que la oferta aumente y el precio disminuya hasta que los beneficios sean nulos; de forma similar, beneficios negativos llevan a que algunas firmas se retiren del mercado disminuyendo la oferta y aumentando el precio hasta el nivel en que los beneficios sean nulos) y que satisfaga el requerimiento tecnológico de que las intensidades del proceso

[4] Arrow, K. y G. Debreu (1954), Existence of an Equilibrium for a Competitive Economy, *Econometrica*, vol. 22, 265-290.

[5] Kakutani, S. (1941), A Generalization of Brouwer's Fixed Point Theorem, *Duke Mathematical Journal*, vol. 8 (3), 457-459.

durante cualquier período no requieran más que los insumos disponibles (es decir, los productos del período anterior). A este comportamiento de la economía lo llama de *equilibrio*, y demuestra que, efectivamente, existe tal tasa de crecimiento; que esta tasa de crecimiento es *igual* a la tasa de interés que satisface la condición de que la tasa de crecimiento de la producción sea exactamente suficiente para cubrir el costo de inversión en insumos (que es una consecuencia de la condición de beneficios cero); que podría haber *muchas* combinaciones producto-precio de equilibrio, a menos que se adicionen hipótesis al modelo; que podría tenerse procesos donde el empleo implica pérdidas financieras pero que estos procesos, en equilibrio, no se utilizan; y, finalmente, que algunas producciones podrían crecer a una tasa mayor que la de equilibrio en algunos períodos, pero que no habrá tasa de crecimiento sostenible mayor que la de equilibrio. Veamos cómo opera esta estructura.

I). El modelo

1. En notación del propio von Neumann, consideremos una economía donde hay n bienes $G_1, G_2, ..., G_n$ que pueden producirse mediante m procesos $P_1, P_2, ..., P_m$, que "para evitar posteriores complicaciones" se asume que tienen *rendimientos constantes a escala*, y que los factores naturales de producción, incluyendo la mano de obra, pueden expandirse en cantidades ilimitadas. Y lo que se pregunta es: *i*) ¿Cuál es la velocidad relativa con la que crece la cantidad de bienes producidos?; *ii*) ¿A qué precios se venderán?; *iii*) ¿Cuál es la tasa de interés ? Para ello, entonces, asume más: el consumo de bienes toma lugar *sólo* a través del proceso de producción que incluye los bienes necesarios de los trabajadores.

2. En cada proceso $P_i(i = 1, 2, ..., m)$ se utilizan cantidades conocidas a_{ij} (expresadas en unidades convenientes) y se producen las cantidades conocidas b_{ij}, de los respectivos bienes $G_j(j = 1, 2, ..., n)$. El proceso, entonces, puede escribirse de la siguiente forma:

$$P_i = \sum_{j=1}^{n} a_{ij}G_j \to \sum_{j=1}^{n} b_{ij}G_j \tag{5}$$

   Estos procesos $P_i(i = 1, 2, ..., m)$ serán utilizados con ciertas intensidades $x_i(i = 1, 2, ..., m)$, lo que significa que, para la producción total, las cantidades de la ecuación (5) deben multiplicarse por x_i. Von Neumann escribe entonces

$$E = \sum_{i=1}^{m} x_i P_i \tag{6}$$

   donde $x_i = 0$ significa que el proceso P_i no será utilizado. Y luego se pregunta por aquellos estados en donde la economía se expande sin cambio de estructura; es decir, donde las proporciones de las intensidades $\frac{x_1}{x_2}$, $\frac{x_2}{x_3}$, ..., $\frac{x_{m-1}}{x_m}$ igualan un factor común α. A éste lo llama el *coeficiente de expansión de la economía.*

3. Las incógnitas del modelo son, entonces,

   i) Las intensidades $x_1, ..., x_m$ de los procesos $P_1, ..., P_m$;

   ii) El *coeficiente de expansión* (o tasa de crecimiento), α, de la economía;

   iii) Los *precios* $y_1, ..., y_n$ de los bienes $G_1, ..., G_n$;

   iv) El *factor de interés* β, donde asume que $\beta = \frac{y_1}{y_2} = \frac{y_2}{y_3} = ... = \frac{y_{n-1}}{y_n}$.

   Obviamente, se deberá asumir

$$x_i \geq 0 \tag{7}$$

$$y_j \geq 0 \tag{8}$$

con al menos alguna x_i, y alguna y_j, estrictamente positivas.

Las ecuaciones económicas son, para $j = 1, 2, ..., n$,

$$\alpha \sum_{i=1}^{m} a_{ij}x_i \leq \sum_{i=1}^{m} b_{ij}x_i \tag{9}$$

ó

$$\alpha A^T X \leq B^T X$$

donde $A = [a_{ij}]_{m\times n}$, $B = [b_{ij}]_{m\times n}$ y $X = [x_1, ..., x_m]^T$ y $\alpha = \frac{x_1}{x_2} = \frac{x_2}{x_3} = ... = \frac{x_{m-1}}{x_m}$; es decir, es imposible consumir de un bien G_j en el proceso total (6) más que lo que está siendo producido. Y para $i = 1, 2, ..., m$,

$$\beta \sum_{j=1}^{n} a_{ij}y_j \geq \sum_{j=1}^{n} b_{ij}y_j \tag{10}$$

ó

$$\beta AY \geq BY$$

donde $Y = [y_1, ..., y_n]^T$ y $\beta = \frac{y_1}{y_2} = ... = \frac{y_{n-1}}{y_n}$; es decir, en "equilibrio" no puede haber beneficio.

Von Neumann, sin embargo, hace las siguientes salvedades:

- Si en (9) se tiene la desigualdad estricta, entonces $y_j = 0$ (9')

es decir, si se consume menos de lo que se produce de algún bien G_j, entonces su precio cae a $y_j = 0$;

- Si en (10) se tiene la desigualdad estricta, entonces $x_i = 0$ (10')

es decir, si P_i da pérdidas entonces no se utiliza, y su intensidad será nula.

Si tenemos en cuenta la condición $\frac{x_1}{x_2} = ... = \frac{x_{m-1}}{x_m} = \alpha$ y la condición $\frac{y_1}{y_2} = ... = \frac{y_{n-1}}{y_n} = \beta$, entonces (9) y (10) conforman un sistema de $m+n$ desigualdades con $m+n$ incógnitas. Pero como éstas no son ecuaciones sino *inecuaciones*, el hecho de que el número de ellas iguale el número de incógnitas, no constituye ninguna garantía de que el sistema pueda resolverse.

Von Neumann entonces prueba utilizando por primera vez en la economía matemática, el teorema de punto fijo de Brouwer, y técnicas de la teoría de programación lineal (Volumen 3 (Optimización y dinámica)) que el sistema $(7)-(10')$ *tiene al menos una solución* $x_1, ..., x_m$*;* $y_1, ..., y_n$*;* α*;* β*. Pero que si se tiene la condición* $a_{ij} + b_{ij} > 0$ *entonces* α *y* β *estarán determinados unívocamente y que, además, (notablemente)* $\alpha = \beta$*; es decir, el factor de interés y el coeficiente de expansión de la economía son iguales y únicamente determinados por los procesos productivos* $P_1, ..., P_m$.

Finalmente, von Neumann hace las tres siguientes observaciones:

i) Que, aunque $\alpha > 0$ por hipótesis, podríamos tener $\alpha \gtreqless 1$; y que aunque uno esperaría $\alpha > 1$, la posibilidad $\alpha \leq 1$ no podría excluirse: los procesos $P_1, ..., P_m$ podrían ser "subproductivos".

ii) Que el *máximo* factor de expansión de la economía es $\alpha = \beta$; y así, podría haber períodos de expansión por debajo del equilibrio.

iii) Que el *mínimo* factor de interés en el que el sistema económico no recibe beneficios es $\beta = \alpha$; y así, podría haber períodos de expansión con tasas de interés superiores a las de equilibrio.

Algo de detalle nos muestra que, de esta manera, el modelo de von Neumann y su característica "dinámica" nos permite combinar tecnologías tipo Leontief distribuyendo los recursos disponibles entre ellas y así evitando las "rigideces" propias de la no-sustitución característica del modelo Leontief. Es decir, en el modelo de von Neumann el conjunto de posibilidades tiene "más esquinas (bordes)" que el modelo de Leontief y, por tanto, está a mitad de camino entre este último y el modelo de conjuntos *"suaves"* de posibilidades con que se aborda actualmente buena parte del análisis de la teoría de la producción (figura 10).

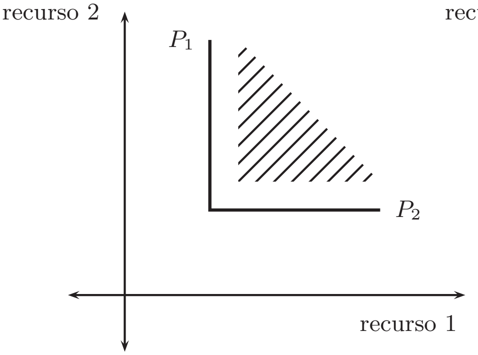


Figura 10-a. Conjunto de posibilidades del modelo de Leontief

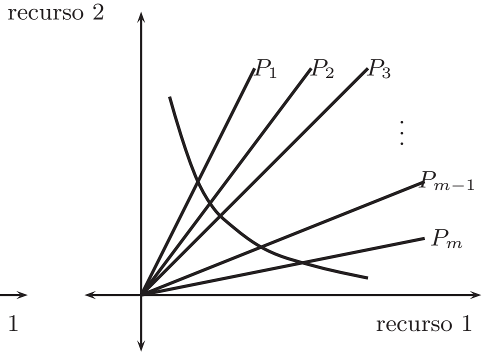


Figura 10-c. Conjunto de posibilidades del modelo de von Neumann

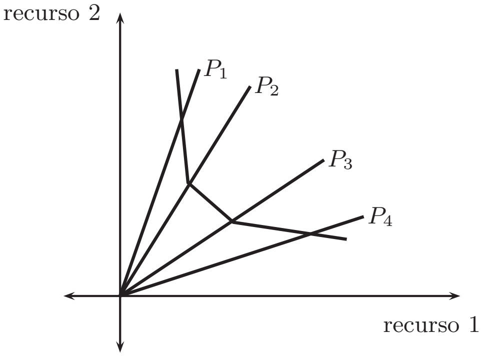


Figura 10-b. Conjunto de posibilidades del modelo von Neumann

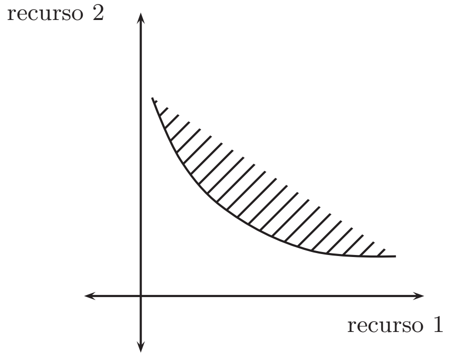


Figura 10-d. Conjunto de posibilidades en modelos "neoclásicos"

Ejemplo 31. (El caso 2×2)
En el caso $m = n = 2$ el modelo de von Neumann se reduce a las siguientes desigualdades:

$$\alpha[a_{11}x_1 + a_{21}x_2] \leq b_{11}x_1 + b_{21}x_2$$

$$\alpha[a_{12}x_1 + a_{22}x_2] \leq b_{12}x_1 + b_{22}x_2$$

$$\beta[a_{11}y_1 + a_{12}y_2] \geq b_{11}y_1 + b_{12}y_2$$

$$\beta[a_{21}y_1 + a_{22}y_2] \geq b_{21}y_1 + b_{22}y_2$$

donde $\dfrac{x_1}{x_2} = \alpha = \dfrac{y_1}{y_2} = \beta$ constante, $x_1, x_2, y_1, y_2 \geq 0$, con al menos alguna x_i y alguna y_j, estrictamente positivas. Y esto nos lleva a un sistema equivalente que es

$$\begin{aligned}
\alpha^2 a_{11} + \alpha a_{21} &\leq \alpha b_{11} + b_{21} \\
\alpha^2 a_{12} + \alpha a_{22} &\leq \alpha b_{12} + b_{22} \\
\alpha^2 a_{11} + \alpha a_{12} &\leq \alpha b_{11} + b_{12} \\
\alpha^2 a_{21} + \alpha a_{22} &\leq \alpha b_{21} + b_{22}
\end{aligned}$$

Ejemplo 32.

a) Supongamos que la tecnología está determinada por $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ y $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$; es decir, $G_1 \to G_2$ y $G_2 \to G_1$ según notación del propio von Neumann. Puesto que $a_{ij}+b_{ij} > 0$ en este caso, entonces la condición $\dfrac{x_1}{x_2} = \alpha = \beta = \dfrac{y_1}{y_2}$ nos lleva al sistema

$$\begin{aligned}
1 \cdot \alpha^2 + 0 \cdot \alpha &\leq 0 \cdot \alpha + 1 \\
0 \cdot \alpha^2 + 1 \cdot \alpha &\leq 1 \cdot \alpha + 0 \\
1 \cdot \alpha^2 + 0 \cdot \alpha &\geq 0 \cdot \alpha + 1
\end{aligned}$$

y esto nos conduce a que $\alpha^2 = 1$, y así $\alpha = \beta = 1$ (crecimiento equilibrado con tasa de interés equilibrada) es la solución única del sistema. ¿Quedan determinados x_1, x_2, y_1, y_2?

b) Supongamos ahora que la tecnología de la economía está determinada por $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ y $B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$; es decir, $G_1 \to 2G_1$ y $G_2 \to 3G_2$.

Observemos que aquí $a_{ij} + b_{ij} \geq 0$ (la desigualdad no es estricta). El sistema de von Neumann nos lleva en este caso a

$$\begin{aligned}
\alpha[x_1 + 0] &\leq 2x_1 \\
\alpha[0 + x_2] &\leq 0 + 3x_2 \\
\beta[y_1 + 0] &\geq 2y_1 + 0 \\
\beta[0 + y_2] &\geq 0 + 3y_2
\end{aligned}$$

o, lo que es lo mismo, a

$$\begin{aligned}
\alpha x_1 \leq 2x_1, &\quad \alpha x_2 \leq 3x_2 \\
\beta y_1 \geq 2y_1, &\quad \beta y_2 \geq 3y_2
\end{aligned}$$

Si $\alpha = 1$, $\beta = 3$; $x_1 = x_2 = 1$; $y_1 = 3$, $y_2 = 1$ el sistema se satisface. Nótese que, sin embargo, $\alpha \neq \beta$. ¿Existirán otras soluciones de este sistema? ▲

En su momento (y todavía) el modelo de von Neumann tuvo un profundo impacto en el desarrollo del pensamiento económico. Como modelo "dinámico" de una economía en expansión, ha sido el padre de muchos modelos de crecimiento; pero también como modelo estático ha tenido influencia. El hecho de que fuera construido como una secuencia de etapas de un solo período, y que cada uno de estos períodos pudiera ser considerado estático (ya que no había oportunidades de ajuste *dentro* de ellos), hizo que pareciera un modelo Walras-Cassel extendido, aunque también era un modelo de *análisis de actividades*. Pero éste sólo aparecería en la literatura casi 20 años después con los trabajos de T. Koopmans (lección 8).

Nota 4. (Sobre John von Neumann)
John von Neumann [1903-1957] hizo importantes contribuciones a la teoría económica de la posguerra a través de dos trabajos seminales: su artículo de (1932, 1946)[6] sobre un modelo de crecimiento multisectorial, y su libro de 1944[7] (con Oskar Morgenstern) sobre teoría de juegos e incertidumbre.

En su artículo de 1932, aparentemente inspirado en su lectura de Wicksell y Cassel, von Neumann dirigió su atención a la teoría del equilibrio general y a la teoría del capital y crecimiento, como acabamos de ver. Entre otros, algunos de los más importantes aportes de este modelo fueron: introducir el concepto de conjuntos de producción en el análisis de actividades, el cual fue utilizado posteriormente por Koopmans y los neo-walrasianos; el sistema de producción descrito en su artículo que luego fue desarrollado aún más por Leontief, Sraffa y los neo-ricardianos; establecer el concepto de crecimiento "balanceado" (*steady-state growth*) empleado posteriormente por Harrod, Solow y Hicks; y la derivación de la *"Regla de Oro"* anticipando los resultados de Allais, Koopmans, Radner y otros. Adicionalmente, fue von Neumann el primero en utilizar la generalización del teorema de punto fijo de Brouwer para probar la existencia de un equilibrio, generalizado posteriormente como el teorema de punto fijo de Kakutani [8] (Volumen 3 (Optimización y dinámica)).

Pero también su libro de 1944 con Oskar Morgestern fue un hito del siglo XX en las ciencias sociales. En él aparecerían fundamentales elementos utilizados

---

[6] Von Neumann, J. (1946), *A Model of General Economic Equilibrium.* In K. Menger, editor, Ergebnisse Eines Mathematischen Kolloquiums, 1935-36. Traducción al inglés de O. Morgenstern (1945).

[7] Von Neumann, J. and O. Morgenstern (1944), *Theory of Games and Economic Behavior*, Princeton: Princeton University Press.

[8] Kakutani, S. (1941), A Generalization of Brouwer's Fixed Point Theorem, *Duke Mathematical Journal*, vol. 8 (3), 457-459.

en economía como, entre otros, la axiomatización de la teoría de la utilidad *per se*, y la axiomatización de la elección bajo incertidumbre (hipótesis de la utilidad esperada) (Volumen 3 (Optimización y dinámica)).

Una anécdota final: John Hicks (premio Nobel en economía 1972 y autor del clásico *Valor y Capital* de 1939) recordaba que en 1933, él y Nicholas Kaldor, el famoso economista húngaro, se encontraron con von Neumann en Budapest, y que éste les enseñó su artículo sobre crecimiento (después de haber sido presentado en Princeton en 1932). De esto, Hicks diría posteriormente: *Por supuesto [que en ese entonces], no entendí absolutamente nada de lo que estaba diciendo.*

Ejercicios complementarios

1) ¿Cuáles de las siguientes funciones $T:\mathbb{R}^2 \longrightarrow \mathbb{R}^2$ son transformaciones lineales?:

a) $T(\,x, y\,) = (\,x - y, 0\,)$ b) $T(\,x, y\,) = (\,x^2 + x + 1, y^2 + y + 1\,)$

2) ¿Será que la *función determinante*

$$\begin{aligned} \det :\ \mathfrak{M}_{n\times n} &\longrightarrow \mathbb{R} \\ A &\longrightarrow \det A \end{aligned}$$

es una transformación lineal?

3) ¿Será que la *función matriz inversa*

$$\begin{aligned} T:\ \mathfrak{M}^*_{n\times n} &\longrightarrow \mathfrak{M}^*_{n\times n} \\ A &\longrightarrow T(A) = A^{-1} \end{aligned}$$

donde $\mathfrak{M}^*_{n\times n}$ es el espacio vectorial de las matrices inversibles $n \times n$, es una transformación lineal? Si la respuesta es negativa, explique, geométricamente, por qué esto es así.

4) a) ¿Será que la función

$$\begin{aligned} T:\ \mathbb{R}^2 \times \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (\,x, y\,) &\longrightarrow x \cdot y \quad \text{(producto interno)} \end{aligned}$$

es una transformación lineal?

b) ¿Y qué acerca de

$$\begin{aligned} T:\ \mathbb{R}^3 \times \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (\,x, y\,) &\longrightarrow x \times y \quad \text{(producto cruz)?} \end{aligned}$$

5) ¿Es $T:\mathfrak{M}_{n\times n} \longrightarrow \mathfrak{M}_{n\times n}$, definida por $T(\,A\,) = AB - BA$ para $B \in \mathfrak{M}_{n\times n}$ fija, una transformación lineal?

6) Encuentre el rango de las transformaciones lineales $T(\,x\,) = Ax$, donde

a) $A = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 18 & 14 \\ -3 & 5 & 36 \end{bmatrix}$ b) $A = \begin{bmatrix} 1 & 4 & 3 \\ 5 & 0 & 9 \\ 1 & -1 & 0 \\ 7 & 2 & 6 \end{bmatrix}$

7) Si T es el operador sobre $\mathbb{R}^4$ definido por

$$T(x, y, z, w) = (x - y + 2z - w, -x + 2z + w, 2x + 2y - z, z - w)$$

a) ¿Cuál es la matriz de T en la base canónica?

b) ¿Cuál es la matriz de T en la base $\{(1, -2, 1, 1), (3, 0, 2, -2), (0, 4, -1, 1), (5, 0, 3, -1)\}$?

c) ¿Cuál es la matriz P que muestra la similaridad entre estas dos matrices?

8) Compruebe el *teorema fundamental del álgebra lineal* para cada una de las siguientes transformaciones lineales:

a) $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por $T(x, y) = (4x + 2y, 6x + 3y)$

b) $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por $T(x, y) = (x - 2y, 3x + 7y)$

c) $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definida por $T(x, y, z) = (5x - y - 6z, -x - y - 4z)$

d) $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ definida por $T(x, y, z) = (-5x, -4y + 2z, 4x - y - z, -x - 7z)$

*9) a) Reflexione sobre la siguiente afirmación:
"Geométricamente, encontrar una solución al sistema lineal $n \times n$, $AX = b$, es hallar el vector de coeficientes X tal que b sea descrito (como combinación lineal con los coeficientes X) en un sistema cartesiano cuyos *ejes* tienen la dirección de las columnas de A."

b) Sabemos (lección 2) que un sistema lineal $m \times n$, $AX = b$, tiene soluciones de la forma $X = z + w$ donde z es una solución al sistema lineal homogéneo $AX = 0$, y w es una solución *particular* del sistema original $AX = b$, pues, recordemos, $A(X - z) = AX - Az = b - 0 = b$, y así $w = X - z$ es una solución particular de $AX = b$. ¿Cuál puede ser una interpretación *geométrica* de esto?

c) ¿Por qué será que la afirmación hecha en la presente lección de que un sistema $AX = 0$, donde A es $n \times n$, tendrá *infinitas* soluciones si $\rho(A) < n$, implica, en particular, que un sistema $AX = 0$, donde A es $m \times n$ con $m < n$, tendrá *infinitas* soluciones? [Indicación: el sistema $m \times n$ $AX = 0$ con $m < n$ puede completarse a un sistema $m \times m$ mediante ecuaciones "redundantes" (es decir, combinaciones lineales de las ecuaciones efectivas)]. ¿Será cierto lo mismo para un sistema $AX = 0$ donde A es $m \times n$ con $m > n$?

10) La figura 11 ilustra cierto proceso productivo conformado por tres procesos lineales P_1, P_2, P_3.

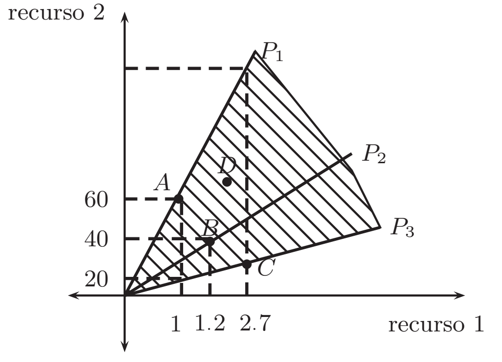


Figura 11. Gráfica del análisis de actividades

a) Describa numéricamente los puntos A, B, C.

b) Describa las *proporciones constantes* entre los recursos para cada uno de los procesos P_1, P_2, P_3.

c) ¿Qué significan los puntos en las regiones *interiores* al cono de la figura 11? Es decir, ¿qué significado tiene dentro del proceso productivo un punto tal como D en la figura?

d) Encuentre tres puntos (uno en cada proceso) que permitan producir la misma cantidad de producto si los puntos A, B, C producen una unidad de producto. Haga esto para varios niveles de producto y conecte linealmente a cada tripla de puntos. ¿Ve usted la formación de curvas de nivel? ¿Qué sucedería si en lugar de tres procesos tuviéramos una cantidad grande de ellos? ¿Cuál sería la forma de las curvas de nivel? ¿Qué significaría tener "infinidad de procesos productivos", en términos del proceso productivo mismo?

11) En cierto modelo von Neumann se tiene que

$$A = [a_{ij}]_{2\times 2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \qquad B = [b_{ij}]_{2\times 2} = \begin{bmatrix} 1.4 & 0 \\ 0 & 1.8 \end{bmatrix}$$

a) Describa la tecnología.

b) Describa con un gráfico el conjunto de posibilidades productivas.

c) Encuentre dos soluciones posibles para $X = (x_1, x_2)$, $Y = (y_1, y_2)$ y las tasas de interés y crecimiento α y β.

$*$12) Algunos autores (por ejemplo, Dorfman, Samuelson y Solow (1958)) aseguran que el modelo de von Neumann incorpora ciertos elementos dinámicos muy interesantes, no vistos antes dentro de la literatura en economía matemática. ¿Podría el lector intentar despejar los puntos exactos del modelo en donde se hallarían estas posibles dinámicas?

Lección 7

Diagonalización en $\mathbb{R}^n$

Introducción

La clase de transformaciones lineales que hemos considerado, a pesar de su aparente simplicidad, es muy general. Sin embargo, puede establecerse que a toda transformación lineal del espacio n-dimensional en sí mismo que satisfaga ciertas condiciones, podemos encontrarle una base en la cual la transformación es muy sencilla: puede describirse mediante una *matriz diagonal*, y ya conocemos las facilidades operacionales de este tipo de matrices: es inmediato calcular la multiplicación de una matriz diagonal por sí misma un número finito de veces; el determinante de una matriz diagonal es simplemente el producto de los términos de la diagonal, etc. Y este proceso, como veremos, está íntimamente conectado con el problema que aquí abordaremos ahora: el cálculo explícito de ciertos números muy importantes conocidos como los *valores propios* de la transformación lineal, y sus correspondientes *vectores propios*. Estudiemos esto en detalle.

1. Valores propios y vectores propios de una transformación lineal

Comenzamos entonces con la definición del concepto central de esta lección: el de *valor propio*.

Definición 1. (Valor propio de una transformación lineal (d'Alembert (1743)))
Sea $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ una transformación lineal. Diremos que $\lambda \in \mathbb{R}$ es un *valor propio (autovalor ó espectro) de la transformación lineal* T, si existe $x \in \mathbb{R}^n, x \neq 0$ tal que

$$T(x) = \lambda x$$

A este x se le denomina un *vector propio* de T (asociado a λ).

Nota 1. (**Los vectores propios como subespacio de** $\mathbb{R}^n$)
Los vectores propios de una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^n$ asociados al valor propio λ, junto con el 0 (cero) de $\mathbb{R}^n$, *conforman un subespacio de* $\mathbb{R}^n$, pues, para $x_1, x_2 \in \mathbb{R}^n,\ k \in \mathbb{R}$,

a) Si $T(x_1) = \lambda x_1$ y $T(x_2) = \lambda x_2$, entonces

$$T(x_1 + x_2) = T(x_1) + T(x_2) = \lambda x_1 + \lambda x_2 = \lambda(x_1 + x_2)$$

y así, $x_1 + x_2$ es un vector propio asociado a λ, si x_1 y x_2 lo son.

b) Si $T(x_1) = \lambda x_1$, entonces

$$T(kx_1) = kT(x_1) = k(\lambda x_1) = \lambda(kx_1)$$

y así, kx_1 es un vector propio asociado a λ, si x_1 lo es.

A este subespacio se le conoce como el *espacio propio de T asociado al valor propio λ.*

Nota 2. (**¿Cómo calcular valores y vectores propios?**)
Si $A = [\, a_{ij}\,]_{n\times n}$ es una matriz que representa a la transformación T en determinada base, entonces λ es un valor propio de T si existe un $x \in \mathbb{R}^n$, $x \neq 0$, tal que $Ax = \lambda x$; es decir, $(A - \lambda I)x = 0$. En forma expandida, esta última ecuación resulta en el sistema homogéneo

$$\begin{aligned}(a_{11} - \lambda)x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0\\ a_{21}x_1 + (a_{22} - \lambda)x_2 + \cdots + a_{2n}x_n &= 0\\ \vdots \qquad \vdots \qquad \vdots\ &\ \vdots\\ a_{n1}x_1 + a_{n2}x_2 + \cdots + (a_{nn} - \lambda)x_n &= 0\end{aligned} \tag{1}$$

donde $x = (x_1, \ldots, x_n)^T$. Ahora: estamos interesados en una *solución no trivial* $x \neq 0$ y sabemos que una condición suficiente y necesaria para que el sistema (1) tenga una solución no trivial es que el determinante de la matriz del sistema sea *nulo*; es decir, que

$$\det \begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} - \lambda & \cdots & a_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{bmatrix} = 0$$

Así, todos los valores propios de la transformación T (representada por A) son las raíces del polinomio $\det(A - \lambda I)$. A este polinomio se le llama *polinomio característico de la transformación* T (término acuñado por Cauchy en 1840). Puesto que por el teorema fundamental del álgebra (volumen 0 (Fundamentos)) todo polinomio de grado n tiene n raíces (que pueden ser complejas y no todas necesariamente diferentes), entonces toda transformación lineal tiene al menos un valor propio (y, a lo sumo, n valores propios diferentes) y, por tanto, al menos un vector propio. Para entender esto último mejor, ilustremos con algunos ejemplos.

Ejemplo 1.
Calculemos los valores y vectores propios de la transformación $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por

$$T(x_1, x_2) = A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \text{donde } A = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}$$

Solución

Resolviendo $\det(A - \lambda I) = 0$, es decir,

$$\det \begin{bmatrix} -5-\lambda & 2 \\ 2 & -2-\lambda \end{bmatrix} = 0$$

se obtiene que

$$\lambda^2 + 7\lambda + 6 = (\lambda + 6)(\lambda + 1) = 0$$

y, por tanto, los valores propios de la transformación T son $\lambda_1 = -1$ y $\lambda_2 = -6$.

Calculando también los correspondientes vectores propios tendremos que:

a) Para $\lambda_1 = -1$, el sistema homogéneo $[A - \lambda_1 I] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ es

$$\begin{aligned} -4x_1 + 2x_2 &= 0 \\ 2x_1 - x_2 &= 0 \end{aligned}$$

y un vector propio es $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ (obsérvese que satisface el sistema anterior). El espacio propio asociado al valor propio $\lambda_1 = -1$ es el espacio generado por el vector propio $(1, 2)$; es decir, $\langle (1, 2) \rangle$.

b) Para $\lambda_2 = -6$, el sistema homogéneo $[A - \lambda_2 I]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ es

$$\begin{aligned} x_1 + 2x_2 &= 0 \\ 2x_1 + 4x_2 &= 0 \end{aligned}$$

y un vector propio es $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ (obsérvese que satisface el sistema anterior). El espacio propio asociado a este valor propio $\lambda_2 = -6$ es $\langle\,(2, -1)\,\rangle$.

Ejemplo 2.
Calculemos los valores y vectores propios de la transformación $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por

$$T(\,x_1, x_2, x_3\,) = A\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \text{donde } A = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$$

Solución
Inicialmente, resolvamos $\det(\,A - \lambda I\,) = 0$; es decir,

$$\det\begin{bmatrix} -2-\lambda & 2 & -3 \\ 2 & 1-\lambda & -6 \\ -1 & -2 & -\lambda \end{bmatrix} = 0$$

De allí se tiene que

$$\lambda^3 + \lambda^2 - 21\lambda - 45 = (\,\lambda - 5\,)(\,\lambda + 3\,)^2 = 0$$

Las raíces del polinomio característico (valores propios) son, entonces, $\lambda_1 = 5$ y $\lambda_2 = \lambda_3 = -3$; y luego calculamos los correspondientes vectores propios:

a) Para $\lambda_1 = 5$, se encuentra el sistema homogéneo $[A - 5I]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, que es, explícitamente,

$$\begin{aligned} -7x_1 + 2x_2 - 3x_3 &= 0 \\ 2x_1 - 4x_2 - 6x_3 &= 0 \\ -x_1 - 2x_2 - 5x_3 &= 0 \end{aligned}$$

y, después de aplicar el método gaussiano, encontramos que su solución es el espacio $\langle\,(1, 2, -1)\,\rangle$; y así, un vector propio asociado es $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$. El espacio propio asociado al valor propio $\lambda_1 = 5$ es, entonces, $\langle\,(1, 2, -1)\,\rangle$.

b) Para $\lambda_2 = \lambda_3 = -3$, el sistema homogéneo asociado es $[A+3I]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} =$

$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$; es decir,

$$\begin{aligned} x_1 + 2x_2 - 3x_3 &= 0 \\ 2x_1 + 4x_2 - 6x_3 &= 0 \\ -x_1 - 2x_2 + 3x_3 &= 0 \end{aligned}$$

y al resolver este sistema se encuentra que los vectores $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ y $\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$ son dos vectores propios *linealmente independientes* asociados al valor propio $\lambda_2 = \lambda_3 = -3$. Claramente, el espacio propio asociado a este valor propio es el espacio $\langle\, (-2, 1, 0), (3, 0, 1)\,\rangle$.

Ejemplo 3. (¿Geométricamente, qué significan los valores propios?) Un buen ejemplo sobre cómo podemos entender los valores propios, es el de un material plástico en el plano xy cuya frontera es la circunferencia $x_1^2 + x_2^2 = 1$, y que se estira de tal modo que un punto $M(x_1, x_2)$ del material original se transforma en el punto $N(y_1, y_2)$ dado por $Ax = y$, donde

$$A = \begin{bmatrix} 9 & 4 \\ 4 & 9 \end{bmatrix}, \qquad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

es decir,

$$\begin{aligned} 9x_1 + 4x_2 &= y_1 \\ 4x_1 + 9x_2 &= y_2 \end{aligned}$$

Encontremos las *direcciones principales*; es decir, las direcciones del vector de posición de M para las que la dirección del vector de posición de N es la misma (o exactamente opuesta) (figura 1). ¿Qué forma asume la circunferencia de la frontera bajo esta deformación?

Solución

Se trata de encontrar vectores x tales que $y = \lambda x$ para algún $\lambda \in \mathbb{R}$. Pero como $y = Ax$, entonces el problema es encontrar λ tal que

$$Ax = \lambda x$$

es decir, un problema de valores propios. La ecuación característica es

$$\begin{vmatrix} 9-\lambda & 4 \\ 4 & 9-\lambda \end{vmatrix} = (9-\lambda)^2 - 16 = 0$$

con soluciones $\lambda_1 = 13$ y $\lambda_2 = 5$. Vectores propios correspondientes son

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ para } \lambda_1 = 13 \; ; \qquad \begin{bmatrix} -1 \\ 1 \end{bmatrix} \text{ para } \lambda_2 = 5$$

Estos vectores forman ángulos de 45º y 135º con la dirección x_1 positiva, y éstas son las direcciones principales (figura 1). *Los valores propios indican que, en las direcciones principales de la superficie, ésta se estira en factores de 13 y 5. Es decir, los valores propios son "coeficientes de dilatación" de los ejes.*

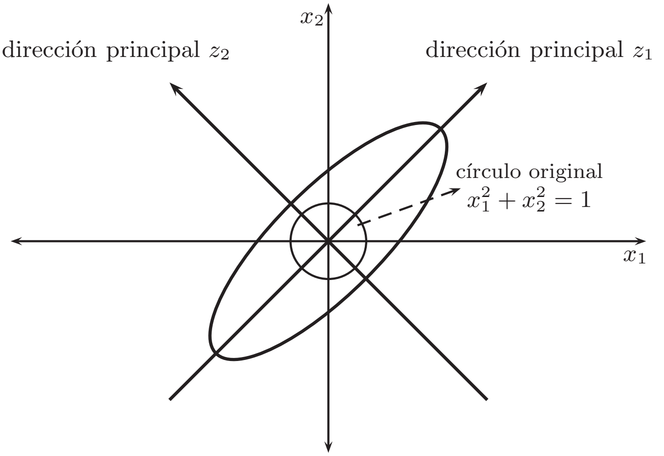


Figura 1

El borde deformado del material plástico es, entonces, una elipse cuya ecuación en los nuevos ejes z_1, z_2, está dada por

$$\frac{z_1^2}{13^2} + \frac{z_2^2}{5^2} = 1$$

Nota 3. (Sobre el concepto de valor propio)
El concepto de *valor propio* apareció en el contexto de las soluciones a sistemas de ecuaciones diferenciales ordinarias, cuando J. L. R. d'Alembert [1717–1783], en su *Traité de Dynamique* de 1743, buscaba describir el movimiento vibratorio de una cuerda con masas atadas a ella en varios puntos. Posteriormente, Jacques Sturm [1803–1855] generalizaría el trabajo de d'Alembert a sistemas de ecuaciones más generales. Hoy en día, el cálculo de valores y vectores propios aparece relacionado con el análisis de *estabilidad* de cuerpos rígidos, con la rotación de éstos, y con el estudio de pequeñas oscilaciones y de resonancias, entre otras aplicaciones en mecánica (física). Sin embargo, cabe anotar aquí que el orden de presentación de los conceptos de esta lección se invierte cuando lo estudiamos desde una perspectiva histórica: primero aparecieron los valores propios en el siglo XVIII en estudios de ecuaciones diferenciales

de la mecánica clásica y de la mecánica celeste; después, estas ecuaciones se presentaron utilizando determinantes, alrededor de la década de 1830; y luego estos determinantes fueron asociados con matrices a finales del siglo XIX. El orden que presentamos en esta lección fue establecido en los libros de texto de matemáticas cerca de la mitad del siglo XX.

Sobre cómo se utilizan algunos de los elementos del álgebra lineal que hemos aprendido aquí, en el estudio de las ecuaciones diferenciales que describen diferentes dinámicas, regresaremos posteriormente en la lección 3 del Volumen 3 (Optimización y dinámica).

Ejercicios 1

1) Encuentre los valores propios *reales* y su vectores propios asociados para las transformaciones lineales $T(x) = Ax$, donde

   a) $A = \begin{bmatrix} 3 & 4 \\ 1 & -1 \end{bmatrix}$ b) $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

   c) $A = \begin{bmatrix} -5 & 2 \\ -8 & 7 \end{bmatrix}$ d) $A = \begin{bmatrix} 2 & -2 & 3 \\ -2 & -1 & 6 \\ 0 & 1 & 3 \end{bmatrix}$

   e) $A = \begin{bmatrix} -1 & 0 & 4 \\ 1 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$ f) $A = \begin{bmatrix} 1 & \frac{1}{4} & 0 \\ 0 & \frac{3}{4} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} \end{bmatrix}$

2) a) ¿Cuáles son los valores propios de una matriz diagonal? En particular, ¿cuáles son los valores propios de la matriz identidad I_n?
   b) ¿Cuáles son los valores propios de la matriz kA para $k \neq 0$, con respecto a los valores propios de la matriz A?
   c) ¿Cuáles son los valores propios de la matriz A^n, para $n \in \mathbb{N}$, con respecto a los valores propios de la matriz A?

*3) ¿Podría el lector explicar el significado geométrico de un valor propio negativo? ¿Y de un valor propio nulo?

2. Diagonalización

Los siguientes teoremas son pasos previos a la posibilidad de “diagonalizar” una matriz cuadrada; es decir, de encontrar una matriz diagonal que sea similar a ella.

Teorema 1.
Si A es similar a una matriz diagonal, entonces los elementos de esta matriz diagonal son los valores propios de A; y así, los valores propios son independientes de la forma en que representemos una transformación lineal.

Demostración

Si $P^{-1}AP = D$ para alguna matriz inversible P, donde D es una matriz diagonal, entonces

$$\det(A - \lambda I) = \det(PDP^{-1} - \lambda I) = \det(P(D - \lambda I)P^{-1}) = \det(D - \lambda I) \quad \blacksquare$$

Teorema 2. ***(Base de vectores propios)***
Si los valores propios de una matriz $n \times n$ son todos números reales diferentes, entonces sus vectores propios correspondientes forman un conjunto linealmente independiente y, por tanto, una base para $\mathbb{R}^n$. La proposición recíproca de este teorema ***no*** *es cierta.*

Demostración

Sean $x_1, x_2, ..., x_n$ vectores propios no nulos correspondientes a los valores propios $\lambda_1, \lambda_2, ..., \lambda_n$, y supongamos que el teorema no es cierto. Sea r el mayor entero tal que el conjunto $\{x_1, x_2, ..., x_r\}$ es linealmente independiente. Entonces existen escalares $k_1, ..., k_{r+1}$, *no todos nulos*, tales que

$$k_1x_1 + \cdots + k_{r+1}x_{r+1} = 0 \tag{1}$$

Si multiplicamos a la izquierda de esta igualdad (1) por la matriz A, obtendremos que

$$k_1\lambda_1x_1 + \cdots + k_{r+1}\lambda_{r+1}x_{r+1} = 0 \tag{2}$$

Y si multiplicamos (1) por λ_{r+1} y restamos (2) de (1) se obtiene que

$$k_1(\lambda_1 - \lambda_{r+1})x_1 + \cdots + k_r(\lambda_r - \lambda_{r+1}) = 0$$

y, por tanto,

$$k_1(\lambda_1 - \lambda_{r+1}) = 0, \ \ldots \ , k_r(\lambda_r - \lambda_{r+1}) = 0$$

Pero como todos los valores propios son distintos, entonces $k_1 = \cdots = k_r = 0$, y así, de (1), también $k_{r+1} = 0$, lo que es una contradicción.

Ahora: que el recíproco de este teorema no es cierto, lo podemos ver en el ejemplo 2 anterior. $\blacksquare$

Ejemplo 4.
En el ejemplo 1, donde los valores propios son $\lambda_1 = -1$ y $\lambda_2 = -6$, la base de vectores propios asociada a la matriz $A = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}$ es $\{(1, 2), (2, -1)\}$.

Ejemplo 5.
Para la matriz $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, los valores propios son $\lambda = 0$ y $\lambda = 2$, con vectores propios $(1, -1)$ y $(1, 1)$, respectivamente, y estos últimos conforman una base para $\mathbb{R}^2$.

Definición 2. (Diagonalización de una matriz)

a) Decimos que una matriz cuadrada A es *diagonalizable* si, y sólo si, existe una matriz invertible P, tal que

$$D = P^{-1}AP$$

es una matriz diagonal. Así, A es diagonalizable si, y sólo si, es similar a una matriz diagonal.

b) Una transformación lineal $T : \mathbb{R}^n \longrightarrow \mathbb{R}$ es diagonalizable si, y sólo si, alguna de sus matrices que la representan en alguna base es una matriz diagonalizable (¿Por qué no depende esta definición de la matriz escogida?).

Teorema 3. ***(Diagonalización de una matriz)***
Una matriz A de tamaño $n \times n$ es diagonalizable si, y sólo si, $\mathbb{R}^n$ tiene una base de vectores propios de A. Cuando es diagonalizable, en la descomposición $D = P^{-1}AP$, las columnas de la matriz P son vectores propios de A.

Demostración
Es una aplicación directa del teorema 6 de la lección 6, y del teorema 2 de esta lección. ■

Ejemplo 6.
La matriz $A = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}$ tiene al vector propio $(1, 2)$ asociado al valor propio $\lambda_1 = -1$ y al vector propio $(2, -1)$ asociado al valor propio $\lambda_2 = -6$. Así, como $\{(1, 2), (2, -1)\}$ es una base para $\mathbb{R}^2$ entonces A es diagonalizable. De hecho, $P = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$, $P^{-1} = -\frac{1}{5}\begin{bmatrix} -1 & -2 \\ -2 & 1 \end{bmatrix}$, y $P^{-1}AP = -\frac{1}{5}$

$$\begin{bmatrix} -1 & -2 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -6 \end{bmatrix}$$

Teorema 4. ***(Condición suficiente para diagonalizar una matriz)***
Si los valores propios de una matriz $n \times n$ son todos números reales diferentes, entonces la matriz es diagonalizable. Sin embargo, la proposición recíproca ***no*** *es cierta.*

Demostración

Este teorema es consecuencia directa de los teoremas 2 y 3. Y que una matriz puede ser diagonalizable aunque no tenga todos sus valores propios reales diferentes, lo podemos ver en el ejemplo 2. ¿Cuál es, allí, la matriz P? ■

Ejemplo 7.

a) La matriz $A = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}$ tiene como vectores propios a $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$ asociado al valor propio $\lambda_1 = 6$ y a $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ asociado al valor propio $\lambda_2 = 1$. Luego, colocando estos vectores propios como columna, obtenemos $P = \begin{bmatrix} 4 & 1 \\ 1 & -1 \end{bmatrix}$ y $P^{-1} = -\frac{1}{5} \begin{bmatrix} -1 & -1 \\ -1 & 4 \end{bmatrix}$ y así

$$P^{-1}AP = -\frac{1}{5} \begin{bmatrix} -1 & -1 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix} = D$$

b) Los valores propios de la matriz $A = \begin{bmatrix} 7.3 & 0.2 & -3.7 \\ -11.5 & 1.0 & 5.5 \\ 17.7 & 1.8 & -9.3 \end{bmatrix}$ son $\lambda_1 = 3$, $\lambda_2 = -4$ y $\lambda_3 = 0$. Vectores propios asociados a estos valores propios son $\begin{bmatrix} -1 \\ 3 \\ -1 \end{bmatrix}$, $\begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$ y $\begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}$, respectivamente. Luego $P = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}$ y $P^{-1} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix}$; y así

$$P^{-1}AP = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

c) La matriz $A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$ tiene vectores propios $x_1 = \begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}$, asociado al valor propio $\lambda_1 = 1$; $x_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$, asociado al valor propio $\lambda_2 = 2$; y, $x_3 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$ asociado al valor propio $\lambda_3 = 2$. Luego $P = \begin{bmatrix} 3 & 2 & 2 \\ -1 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$ y $P^{-1} = \begin{bmatrix} -1 & 2 & 2 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix}$; y así

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

d) La matriz $A = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$ tiene como vectores propios a $x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ asociado al valor propio $\lambda_1 = 0$; a $x_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ asociado al valor propio $\lambda_2 = 1$; y, a $x_3 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ asociado al valor propio $\lambda_3 = 3$. Luego $P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{bmatrix}$ y $P^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} \end{bmatrix}$; y de esta forma,

$$P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

e) La matriz $A = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$ tiene vectores propios $x_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ asociado al valor propio $\lambda_1 = 5$; $x_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ asociado al valor propio $\lambda_2 = -3$; y $x_3 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$ asociado al valor propio $\lambda_3 = -3$. Luego

$$P = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{y} \quad P^{-1} = \begin{bmatrix} 1 & 2 & -3 \\ -2 & 4 & 6 \\ 1 & 2 & 5 \end{bmatrix}; \text{ y así,}$$

$$P^{-1}AP = \begin{bmatrix} 5 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

f) Sin embargo, no todas las matrices poseen n vectores propios linealmente independientes y, por tanto, *no todas las matrices son diagonalizables.* El ejemplo estándar es la matriz

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

cuyos valores propios son $\lambda_1 = \lambda_2 = 0$, pues

$$\det(A - \lambda I) = \det \begin{bmatrix} -\lambda & 1 \\ 0 & -\lambda \end{bmatrix} = \lambda^2$$

y aunque $\lambda = 0$ es un valor propio de multiplicidad 2, su espacio propio *sólo tiene dimensión* 1:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \text{si, y sólo si, } x_2 = 0;$$

es decir, si, y sólo si,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

El espacio propio es el eje x_1 en el plano x_1x_2. ▲

En definitiva, para decidir cuándo una transformación lineal es diagonalizable requerimos de un teorema general que no demostraremos aquí pues nos obligaría a disgresiones más allá del alcance de este texto[1]. Es el siguiente:

Teorema 5. ***(Condición general para diagonalizar una matriz)***
Sea $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ una transformación lineal; sean $\lambda_1, \ldots, \lambda_k$ $(k \leq n)$ los valores propios (distintos) de T y $W_1, ..., W_k$ los espacios de vectores propios

[1] Ver, por ejemplo, Hoffman, Kenneth y Ray Kunze (1971), *Linear Algebra*, Prentice Hall.

asociados a los correspondientes vectores propios. Entonces, T es diagonalizable si, y sólo si, el polinomio característico para T se puede escribir de la forma

$$P(\lambda) = (\lambda - \lambda_1)^{d_1} \cdots (\lambda - \lambda_k)^{d_k}$$

donde $d_i = \dim W_i$ para $i = 1, 2, \ldots, k$, y además

$$\dim W_i + \cdots + \dim W_k = n$$

Ejemplo 8.
Consideremos la transformación lineal $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ dada por la matriz

$$A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$$

Por el ejemplo 7 literal c), sabemos que la matriz A es diagonalizable. Observemos que su polinomio característico es $P(\lambda) = -(\lambda - 1)(\lambda - 2)^2$; el espacio propio W_1 asociado al valor propio $\lambda_1 = 1$ es $\langle (3, -1, 3) \rangle$ y su dimensión es 1; y el espacio propio W_2 asociado al valor propio $\lambda_2 = 2$ es $\langle (2, 1, 0), (2, 0, 1) \rangle$ y su dimensión es 2. Además, $\dim \mathbb{R}^3 = \dim W_1 + \dim W_2$. Por tanto, el polinomio característico asociado a A puede escribirse en la forma dada por el teorema 5.

Ejemplo 9.
Consideremos ahora la transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^3$ determinada por la matriz

$$A = \begin{bmatrix} -1 & 1 & -2 \\ 0 & -1 & 4 \\ 0 & 0 & 1 \end{bmatrix}$$

Aquí, el polinomio característico de A es $p(\lambda) = (\lambda + 1)^2(\lambda - 1)$. El espacio propio W_1 asociado al valor propio $\lambda_1 = -1$ es $\langle(1, 0, 0)\rangle, \langle(0, 1, 0)\rangle$; y el espacio propio W_2 asociado al valor propio $\lambda_2 = 1$ es $\langle(0, 2, 1)\rangle$ como fácilmente puede verificarse. Así, $\dim W_1 + \dim W_2 = 3$ y, por el teorema 5, T es diagonalizable en la matriz

$$D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Ejemplo 10.
La transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^3$ determinada por la matriz

$$A = \begin{bmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{bmatrix}$$

tiene como polinomio característico $p(\lambda) = -\lambda(\lambda+2)(\lambda+3)$. El espacio propio W_1 asociado a $\lambda = 0$ es $\langle(0,1,-1)\rangle$; el espacio propio W_2 asociado a $\lambda = -2$ es $\langle(-2,1,0)\rangle$; y, finalmente, el espacio propio W_3 asociado a $\lambda = -3$ es $\langle(1,0,-1)\rangle$. Así, $\dim W_1 + \dim W_2 + \dim W_3 = 3$ y, por tanto, T es diagonalizable en la matriz

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

Ejercicios 2

1) Pruebe que la matriz

$$A = \begin{bmatrix} -17 & 18 & -6 \\ -18 & 19 & -6 \\ -9 & 9 & -2 \end{bmatrix}$$

es diagonalizable a través de la matriz

$$P = \begin{bmatrix} 2 & 1 & -1 \\ 2 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

2) Pruebe que la transformación representada (en cierta base) por la matriz

$$A = \begin{bmatrix} 9 & 2 & -4 \\ -8 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

es diagonalizable.

3) ¿Es

$$A = \begin{bmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ -10 & -5 & -3 \end{bmatrix}$$

una matriz diagonalizable?

4) Pruebe que las matrices

a) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 2 & -4 \\ 0 & -1 & 6 \\ 0 & -1 & 4 \end{bmatrix}$

no son diagonalizables.

5) Pruebe que las siguientes matrices son diagonalizables:

a) $\begin{bmatrix} \cos\phi & -\operatorname{sen}\phi \\ \operatorname{sen}\phi & \cos\phi \end{bmatrix}$ b) $\begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{2}}{2} & \frac{1}{2} \end{bmatrix}$

c) $\begin{bmatrix} 1 & 0 & 0 & -5 \\ 0 & -1 & 0 & 6 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$ d) $\begin{bmatrix} 3 & -2 \\ 2 & 1 \end{bmatrix}$

*6) a) Muestre que el *determinante* de una matriz cuadrada es el *producto* de sus valores propios.

b) Muestre que la *traza* de una matriz cuadrada es la *suma* de sus valores propios.

c) Muestre que el *rango* de una matriz cuadrada es el *número* de sus valores propios *distintos de cero*.

d) Muestre que los *valores propios* de una matriz idempotente ($A^2 = A$) son 1 ó 0. ¿Cuál es la intuición geométrica de esto?

7) Pruebe que una matriz cuadrada es invertible si, y sólo si, todos sus valores propios son distintos de cero[2]; y, además, que los valores propios de su matriz inversa son los recíprocos de los valores propios de la matriz. [Indicación: utilice el literal a) del ejercicio 6 anterior].

8) Pruebe que si A y B son matrices cuadradas y alguna de las dos es inversible, entonces AB y BA tienen sus valores propios iguales.

3. Diagonalización de matrices simétricas: el teorema espectral

Recordemos que una matriz cuadrada A, es simétrica si $A^T = A$. Estas matrices, que han estado por más de cien años en el centro de las herramientas analíticas de la mecánica clásica y celeste, constituyen, además, uno de los tipos más importantes de matrices diagonalizables, dado que todos sus valores propios son números reales. Y, más interesante aún, *las matrices simétricas son diagonalizables vía matrices ortogonales.*

[2] Observe el lector que si el sistema $AX = 0$, donde A es $n \times n$, tiene un valor propio nulo, significaría que al menos una ecuación "desaparece" del sistema, y así el nuevo tendría infinitas soluciones, pues el número de ecuaciones efectivas sería menor que el número de incógnitas.

Teorema 6. ***(Teorema espectral para matrices simétricas (Cauchy (1826)))***

a) Todos los vectores propios de una matriz simétrica son reales.

b) Si A es una matriz simétrica entonces $\mathbb{R}^n$ tiene una base ortonormal conformada por vectores propios de A y, por tanto, es diagonalizable mediante una matriz ortogonal.

Demostración

(Esta demostración requiere algún conocimiento de números complejos. En el ejercicio 19 de los ejercicios complementarios, al final de esta lección, damos algunos elementos para su prueba).

Ejemplo 11.

Diagonalizar la matriz simétrica

$$A = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 5 & 4 \\ -4 & 4 & 3 \end{bmatrix}$$

Solución

Sabemos, por el teorema 6, que A es diagonalizable. Con el fin de encontrar esta diagonalización, inicialmente determinemos los valores propios calculando las raíces de su polinomio característico; es decir,

$$\begin{vmatrix} 1-\lambda & 0 & -4 \\ 0 & 5-\lambda & 4 \\ -4 & 4 & 3-\lambda \end{vmatrix} = 0$$

Así obtenemos que los valores para λ son 3, -3 , 9. En el siguiente paso encontramos un vector propio asociado a cada uno de estos valores propios. Por ejemplo, para calcular el vector propio $x = (\,x_1, x_2, x_3\,)^T$ correspondiente al valor propio $\lambda = 3$, resolvemos la ecuación matricial

$$\begin{bmatrix} 1 & 0 & -4 \\ 0 & 5 & 4 \\ -4 & 4 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 3 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

lo cual es equivalente a resolver el sistema homogéneo

$$\begin{array}{rcrcrcl} x_1 & & & + & 2x_3 & = & 0 \\ & & x_2 & + & 2x_3 & = & 0 \\ -x_1 & + & x_2 & & & = & 0 \end{array}$$

Obtenemos de aquí la solución general $(-2x_3, -2x_3, x_3)$. Haciendo $x_3 = 1$, obtenemos $(-2, -2, 1)$ y normalizando este vector obtenemos $\frac{1}{3}(-2, -2, 1)$. Mediante un procedimiento similar, podemos obtener los vectores propios asociados con los valores propios -3 y 9; éstos son $\frac{2}{3}(1, -\frac{1}{2}, 1)$ y $\frac{2}{3}(-\frac{1}{2}, 1, 1)$. Por tanto, tenemos que las columnas de P son precisamente estos vectores propios de A; es decir, la matriz ortogonal es

$$P = \frac{2}{3}\begin{bmatrix} -1 & 1 & -\frac{1}{2} \\ -1 & -\frac{1}{2} & 1 \\ \frac{1}{2} & 1 & 1 \end{bmatrix}$$

Podemos comprobar ahora que, efectivamente,

$$\frac{2}{3}\begin{bmatrix} -1 & -1 & \frac{1}{2} \\ 1 & -\frac{1}{2} & 1 \\ -\frac{1}{2} & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -4 \\ 0 & 5 & 4 \\ -4 & 4 & 3 \end{bmatrix} \frac{2}{3}\begin{bmatrix} -1 & 1 & -\frac{1}{2} \\ -1 & -\frac{1}{2} & 1 \\ \frac{1}{2} & 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

Ejemplo 12.
Encontremos una matriz ortogonal P que diagonalice la matriz

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}$$

Solución
Los valores propios de A son $\lambda = 1, 1, -2$. Una base del espacio propio de $\lambda = 1$ es $\{(1, 0, 1), (1, 2, -1)\}$; y una base para el espacio propio del valor $\lambda = -2$ es $\{(-1, 1, 1)\}$. La matriz ortogonal P tal que $P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$ es

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

como el lector puede comprobar fácilmente.

Teorema 7. ***(Y, finalmente, ¿qué aprendimos sobre los sistemas homogéneos?)***
Sea A una matriz $n \times n$; entonces las ocho afirmaciones siguientes son equivalentes:

a) $AX = 0$ *tiene solución única* $X = 0$.

b) A *es invertible.*

c) $\det A \neq 0$.

d) *Las columnas de* A *forman una base para* $\mathbb{R}^n$. *También las filas de* A *forman una base para* $\mathbb{R}^n$.

e) *La dimensión del espacio de soluciones del sistema homogéneo* $AX = 0$ *es igual a cero.*

f) $\rho(\,A\,) = n$.

g) *La transformación* $T(x) = AX$ *es un isomorfismo de* $\mathbb{R}^n$ *en sí mismo.*

h) A *no tiene valores propios nulos.*

Ejercicios 3

1) En cada una de las siguientes matrices simétricas:

a) $\begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}$ b) $\begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix}$

c) $\begin{bmatrix} 1 & -1 & 0 \\ -1 & 4 & 3 \\ 0 & 3 & 1 \end{bmatrix}$ d) $\begin{bmatrix} 1 & -1 & 0 \\ -1 & 4 & -3 \\ 0 & -3 & 5 \end{bmatrix}$

i) Encuentre los valores propios y una base de vectores propios.

ii) Diagonalice mediante una matriz ortogonal.

4. Formas cuadráticas

Las formas cuadráticas ocurren muy comúnmente en matemáticas y en sus aplicaciones: en la teoría de números y en la cristalografía; en la geometría analítica (en donde representan ecuaciones de curvas de segundo orden); en mecánica (física) (en donde aparecen representando la energía cinética de un sistema en términos de las componentes de las velocidades); incluso en el análisis matemático se hacen necesarias (allí aparecen en el estudio de funciones de varias variables en conexión con el problema de estimar los máximos y los mínimos de una función).

El estudio de las formas cuadráticas se ha centrado, esencialmente, en la investigación del *problema de la equivalencia de sus "formas" bajo transformaciones lineales*. Se dice que dos formas cuadráticas son (linealmente) *equivalentes* si una de ellas puede transformarse en la otra mediante una transformación lineal. En ocasiones, este proceso de simplificación permite que emerjan las características esenciales de la forma cuadrática. Este es el *problema de reducción* sobre el cual nos concentraremos en esta sección.

Definición 3. (Forma cuadrática (A. Cayley (1853)))
Una *forma cuadrática* es un polinomio de grado 2 en varias variables; es decir, es un polinomio de la forma

$$\begin{aligned} Q(x_1, x_2, \ldots, x_n) = {} & a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n + \\ & a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n + \cdots + \\ & a_{n1}x_nx_1 + a_{n2}x_2 + \cdots + a_{nn}x_n^2 \end{aligned} \tag{1}$$

Ejemplo 13.
Las siguientes son formas cuadráticas:

a) $Q(x, y) = x^2 + y^2$ b) $Q(x, y) = \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2}$

c) $Q(x, y) = \dfrac{x^2}{a^2} - \dfrac{y^2}{b^2}$ d) $Q(x, y) = ax^2 + bxy + cy^2$

La razón por la cual estudiamos en este punto a estos polinomios específicos es porque una forma cuadrática puede representarse, convenientemente, en notación matricial, de la siguiente manera:

$$Q(x_1, x_2, \ldots, x_n) = [\,x_1, x_2, \ldots, x_n\,] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = XAX^T$$

donde $A = [a_{ij}]$ es la matriz $n \times n$ conformada por los coeficientes que aparecen en la ecuación (1) anterior y $X = (x_1, x_2, ..., x_n)$. Observemos que podemos asumir que A es *simétrica* ya que los pares de coeficientes similares $a_{12}x_1x_2$ y $a_{21}x_2x_1$, etc., pueden escribirse con coeficientes iguales de tal forma que cada uno de ellos sea la mitad del coeficiente del producto correspondiente de las variables.

La idea central para resolver el problema de reducción es el de *encontrar los ejes principales* de la forma cuadrática que, a su vez, está relacionado con la diagonalización de la matriz A de coeficientes. Veamos cómo.

Sea $Q = XAX^T$ una forma cuadrática, donde A es una matriz *simétrica* $n \times n$. Puesto que A es diagonalizable mediante una matriz ortogonal P, se tiene que

$D = P^{-1}AP$ es diagonal. Así,

$$\begin{aligned} Q &= XAX^T = X(\,PDP^{-1}\,)X^T \\ &= X(\,PDP^T\,)X^T \quad (\text{pues } P^{-1} = P^T) \\ &= YDY^T \\ &= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \end{aligned}$$

donde $Y = XP$; y $\lambda_1, \lambda_2, \ldots, \lambda_n$ son los valores propios de A. Esto nos lleva al siguiente resultado:

Teorema 8. ***(Reducción a ejes principales)***
Toda forma cuadrática $Q = XAX^T$ con A una matriz simétrica $n \times n$, puede reducirse a la forma

$$Q' = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

donde $\lambda_1, \lambda_2, \ldots, \lambda_n$ son valores propios de A. A los vectores propios asociados a $\lambda_1, \lambda_2, \ldots, \lambda_n$ se les llama los ejes principales de la forma cuadrática Q. Es claro entonces que $y_1, y_2, \ldots, y_n$ son nuevas coordenadas cartesianas sobre las cuales se representa a la misma forma cuadrática.

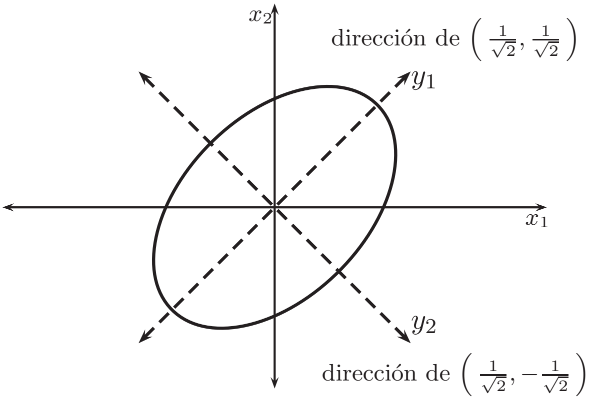


Figura 2. Elipse $2x_1 - 2x_1x_2 + 2x_2^2 = 1$ (ejes viejos)
ó $y_1^2 + 3y_2^2 = 1$ (ejes nuevos)

Ejemplo 14.
Para el caso $n = 2$, la ecuación $Q(\,x_1, x_2\,) = a_{11}x_1^2 + a_{12}x_1x_2 + a_{21}x_2x_1 + a_{22}x_2^2 = 1$ representa una cónica en el plano x_1, x_2. En términos de las nuevas coordenadas y_1, y_2 esta ecuación es

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 = 1;$$

y representa una elipse o una hipérbola. Es decir, toda ecuación de segundo grado de la forma $Ax^2 + Bxy + Cy^2 = 1$ se puede reducir a la forma estándar $A'x^2 + C'y^2 = 1$ mediante una rotación conveniente de ejes en el plano XY. Por ejemplo, la elipse $2x_1^2 - 2x_1x_2 + 2x_2^2 = 1$ puede escribirse como

$$[\, x_1, x_2 \,] \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 1$$

Los valores propios de $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ son $\lambda_1 = 1$ y $\lambda_2 = 3$, y sus vectores propios asociados son $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$ y $\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right)$, respectivamente, que es hacia donde se dirigen los nuevos ejes (figura 2). La nueva ecuación es la elipse $y_1^2 + 3y_2^2 = 1$.

Ejemplo 15.
La forma cuadrática $Q = x_1^2 + 2x_2^2 + 3x_3^2 + 4x_1x_2 + 4x_2x_3$, tiene la forma XAX^T para

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 3 \end{bmatrix}; \quad \text{y} \quad X = (\, x_1, x_2, x_3 \,)$$

Los valores propios son $\lambda_1 = 2$, $\lambda_2 = 5$ y $\lambda_3 = -1$ y sus correspondientes vectores propios normalizados son $\left(\frac{2}{3}, \frac{1}{3}, -\frac{2}{3} \right)$, $\left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3} \right)$ y $\left(\frac{2}{3}, -\frac{2}{3}, \frac{1}{3} \right)$. Por tanto,

$$P^TAP = P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{para} \quad P = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

Luego si $Y = (\, y_1, y_2, y_3 \,) = XP$, entonces la forma se reduce a la cuadrática en tres variables $2y_1^2 + 5y_2^2 - y_3^2 = 1$ que es un hiperboloide de dos hojas (figura 3).

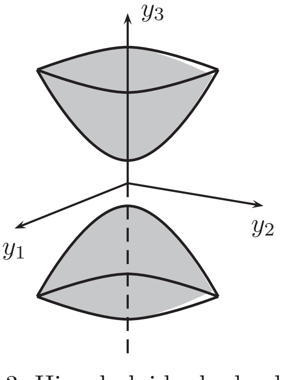


Figura 3. Hiperboloide de dos hojas

En la reducción de una forma cuadrática a una forma canónica siempre existe algo de arbitrariedad en la elección de las variables de transformación que permiten tal reducción. La arbitrariedad aparece, por ejemplo, al elegir el tipo de transformaciones que aplicamos. A pesar de esto, toda forma cuadrática canónica posee algunas características propias que permanecen invariantes bajo cualquier tipo de transformación. Son de especial interés, en las aplicaciones, aquellas formas cuadráticas que bajo la reducción a la forma canónica resultan con *todos los coeficientes cuadráticos positivos*; a tales formas se les llama *formas cuadráticas definidas positivas* y éstas, como veremos, se caracterizan por la propiedad de que sus valores propios son siempre positivos. De manera similar, las *formas cuadráticas definidas negativas* son aquéllas en las que todos sus coeficientes cuadráticos son negativos y, por tanto, sus valores propios son todos negativos. Enseguida estudiamos esto con más detalle.

Definición 4. (Formas cuadráticas definidas positivas (y negativas))

a) Diremos que una forma cuadrática $Q = XAX^T$, con A simétrica, es *definida positiva* si, y sólo si, $XAX^T > 0$ para todo $X \neq 0$. Y diremos que Q es *definida negativa* si $-Q$ es definida positiva.

b) Diremos, además, que $Q = XAX^T$ es *semidefinida positiva* si $XAX^T \geq 0$ para todo X. Además, diremos que Q es *semidefinida negativa* si $-Q$ es semidefinida positiva.

Se ve claro que esta definición se encuentra ligada a conceptos puramente geométricos del espacio tridimensional cuando entendemos que los paraboloides hacia abajo y hacia arriba que estudiamos en la geometría analítica ordinaria, no son más que formas cuadráticas definidas positivas y negativas, respectivamente. Esto lo corroboramos en el siguiente ejemplo, en donde colocamos condiciones precisas sobre los coeficientes de la ecuación cuadrática para que satisfaga las condiciones de positividad o negatividad.

Ejemplo 16.

a) $Q = ax^2 + 2bxy + cy^2 = [\,x, y\,]\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = XAX^T$ es *definida positiva* si, y sólo si,

$$\text{a)}\quad a > 0 \qquad\qquad \text{b)}\quad \det A = ac - b^2 > 0$$

pues $ax^2 + 2bxy + cy^2 = a(\,x + \dfrac{b}{a}y\,)^2 + (\dfrac{ac - b^2}{a}\,)y^2$

b) Y, por tanto, es *definida negativa* si, y sólo si,

$$\text{a)}\quad a < 0 \qquad\qquad \text{b)}\quad \det A = ac - b^2 > 0$$

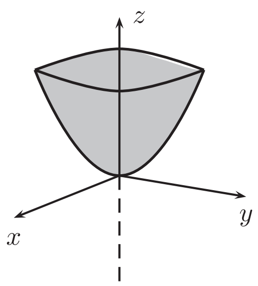


Figura 4. Forma cuadrática definida positiva

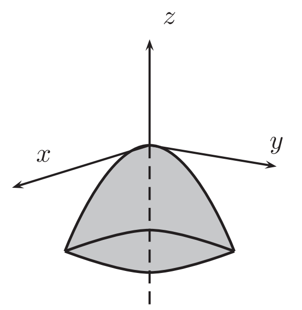


Figura 5. Forma cuadrática definida negativa

Ahora: recordando que, según afirmamos al comienzo de nuestra discusión, toda forma cuadrática $Q = XAX^T$ puede reducirse a la forma $Q' = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$, podemos entonces afirmar lo siguiente:

Teorema 9. ***(Una caracterización de las formas cuadráticas)***

a) *$Q = XAX^T$ es definida positiva si, y sólo si, todos los valores propios de A son estrictamente positivos.*

b) *$Q = XAX^T$ es definida negativa si, y sólo si, todos los valores propios de A son estrictamente negativos.*

Demostración

a) Sea λ cualquier valor propio de A y $X \neq 0$ un vector propio asociado; es decir, $AX^T = \lambda X^T$. Pre-multiplicando esta igualdad por X, obtenemos que $XAX^T = \lambda XX^T = \lambda \|X\|^2$. Así $\lambda = \dfrac{XAX^T}{\|X\|^2}$ y, por tanto, $XAX^T > 0$ si, y sólo si, $\lambda > 0$.

b) Es similar. ■

Ejemplo 17.
Determinemos si la forma cuadrática XAX^T es definida positiva cuando

$$A = \begin{bmatrix} 5 & 9 \\ 1 & 5 \end{bmatrix}$$

Solución
La ecuación característica de A es $P(\lambda) = \lambda^2 - 10\lambda + 16 = 0$ y así, los valores propios son $\lambda_1 = 8$ y $\lambda_2 = 2$. Por el teorema 8, la forma cuadrática es definida positiva. ¿Podría el lector describir la forma cuadrática explícitamente?

Ejemplo 18.
Determinemos si la forma cuadrática XAX^T es definida positiva, donde

$$A = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$$

Solución

Los valores propios para A son 5, -3 y -3. Como dos de estos valores propios son negativos y uno positivo, podemos asegurar que la forma cuadrática no es definida positiva ni definida negativa. ¿Podría el lector escribir la forma cuadrática explícitamente? ▲

Regresando a nuestra discusión general, y entendiendo que los valores propios de una matriz simétrica y los determinantes de sus cofactores están íntimamente relacionados, no debería sorprendernos la siguiente caracterización alterna para las formas cuadráticas:

Teorema 10. ***(Otra caracterización de las formas cuadráticas)***

a) *La forma cuadrática $Q = XAX^T$ es definida positiva si, y sólo si, las submatrices A_k, $k = 1, 2, \ldots, n$, donde*

$$A_1 = a_{11}, \qquad A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix},$$

$$A_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{32} & a_{33} \end{bmatrix}, ..., A_n = A$$

tienen determinantes positivos.

b) *La forma cuadrática $Q = XAX^T$ es definida negativa si, y sólo si, las submatrices A_k, $k = 1, 2, \ldots, n$ satisfacen $|A_1| < 0$, $|A_2| > 0$, $|A_3| < 0$, $\ldots$*

Demostración

a) Estudiaremos sólo el caso $n = 2$; el caso general es similar aunque involucra mayor cantidad de operaciones.

Notemos que de la ecuación característica de A

$$\det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{12} & a_{22} - \lambda \end{vmatrix} = 0$$

se obtiene que

$$\lambda^2 - (\text{traza}\, A)\lambda + \det A = 0$$

Así, si λ_1 y λ_2 son los valores propios de A entonces

$$\lambda_1 + \lambda_2 = \text{traza}\, A = a_{11} + a_{22}\,;\ \lambda_1\lambda_2 = \det A = a_{11}a_{22} - (a_{12})^2 \qquad (1)$$

Aplicando el teorema 8, XAX^T es definida positiva si, y sólo si, λ_1 y λ_2 son positivos; y esto, por (1) arriba, implica que $\det A = a_{11}a_{22} - (a_{12})^2 > 0$ y traza$A = a_{11} + a_{12} > 0$; y así, $a_{11}a_{22} > (a_{12})^2$ y $a_{11} + a_{22} > 0$ implican que, también, $a_{11} > 0$ y $a_{22} > 0$. De manera similar, si $a_{11} > 0$ y $\det A = a_{11}a_{22} - (a_{12})^2 > 0$, entonces, de (1) arriba, λ_1 y λ_2 tienen el mismo signo y, por tanto, ambos son positivos.

b) Esta parte es aplicación directa del hecho de que $\det(-A) = (-1)^n \det A$. ■

Ejemplo 19.
Determinemos si la siguiente matriz es definida positiva, aplicando el teorema 9:

$$A = \begin{bmatrix} 7 & 2 & 4 \\ 2 & 2 & 1 \\ 4 & 1 & 3 \end{bmatrix}$$

Solución

Observemos que

$$|\,A_1\,| = 7 \quad ; \quad |\,A_2\,| = \begin{vmatrix} 7 & 2 \\ 2 & 2 \end{vmatrix} = 10 \quad ; \quad |\,A_3\,| = \begin{vmatrix} 7 & 2 & 4 \\ 2 & 2 & 1 \\ 4 & 1 & 3 \end{vmatrix} = 7$$

Como los determinantes de las submatrices de A son todos positivos, la forma cuadrática XAX^T es definida positiva. De hecho, los valores propios de A son positivos: 0.486, 1.428, y 10.085

Ejemplo 20.
Establezcamos si la siguiente matriz es definida positiva, observando el signo de los determinantes de las submatrices:

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 5 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

Solución

En este caso,

$$|\,A_1\,| = 2 \quad ; \quad |\,A_2\,| = \begin{vmatrix} 2 & 1 \\ 1 & 5 \end{vmatrix} = 9 \quad ; \quad |\,A_3\,| = \begin{vmatrix} 2 & 1 & 0 \\ 1 & 5 & 2 \\ 5 & 2 & 1 \end{vmatrix} = 1 \ ;$$

y así, todos los determinantes de las submatrices de A son positivos y, por tanto, también la forma cuadrática XAX^T es definida positiva. De hecho, los valores propios de la matriz A son positivos: 0.0885, 1.8705, 6.041.

Teorema 11.

a) $Q = XAX^T$ *es semidefinida positiva si, y sólo si, todos los valores propios de* A *son positivos o ceros.*

b) $Q = XAX^T$ *es semidefinida negativa si, y sólo si, todos los valores propios de* A *son negativos o ceros.*

Demostración

Es similar a la del teorema 8. ■

Ejemplo 21.

La forma cuadrática $Q = XAX^T$ donde

$$A = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

es semidefinida positiva pues sus valores propios son $\lambda_1 = 0$, $\lambda_2 = 1$ y $\lambda = 3$ como el lector puede comprobar.

Ejemplo 22.

La forma cuadrática $Q = XAX^T$ donde

$$A = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 5 & 4 \\ -4 & 4 & 3 \end{bmatrix}$$

no es ni semidefinida positiva ni semidefinida negativa pues sus valores propios son $\lambda_1 = 3$, $\lambda_2 = -3$ y $\lambda_3 = 9$.

Ejercicios 4

1) Determine si cada una de las siguientes formas cuadráticas es definida (o semidefinida) positiva o negativa, o ninguna de estas:

   a) $Q(x_1, x_2) = 4x_1^2 - x_2^2$
   b) $Q(x_1, x_2) = 3x_1^2 + 7x_1x_2 - 5x_2^2$
   c) $Q(x_1, x_2) = 6x_1x_2 - 9x_2^2 - x_1^2$
   d) $Q(x_1, x_2) = 4x_1^2 + 8x_1x_2 + 5x_2^2$

2) Determine el tipo de cónica que representan las siguientes formas cuadráticas dadas y transfórmelas a los ejes principales:

   a) $x_1^2 + 14x_1x_2 - 6x_2^2 = 5$
   b) $3x_1^2 + 4\sqrt{3}x_1x_2 + 4x_2^2 = 5$
   c) $3x_1^2 + 4\sqrt{3}x_1x_2 + 7x_2^2 = 9$
   d) $6x_1^2 + 16x_1x_2 - 6x_2^2 = 10$

   ¿Cuáles de ellas son definidas positivas?

3) Reduzca la elipse $3x_1^2 - 2\sqrt{2}x_1x_2 + 2x_2^2 = 1$ a su forma canónica $\dfrac{u^2}{a^2} + \dfrac{v^2}{b^2} = 1$, calculando los valores propios y sus vectores propios correspondientes. Dibújela. ¿Es esta forma cuadrática definida positiva?

4) Decida si la forma cuadrática $Q = XAX^T$ donde

$$A = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 1 & 7 \\ 0 & 7 & 3 \end{bmatrix}$$

   es definida positiva, definida negativa, semidefinida positiva, semidefinida negativa, o ninguna de las anteriores.

5. Breve nota sobre la diagonalización en bloques de Jordan

Aunque ya hemos observado que no toda matriz es diagonalizable, el problema de llevar la matriz de una transformación lineal a una forma "más simple" (en donde sean fácilmente identificables sus valores propios) es un tanto complicado si se presentan valores propios para los cuales la dimensión de su espacio propio es menor que su multiplicidad algebraica dentro del polinomio característico. El siguiente esquema, conocido como *descomposición en bloques de Jordan*, es una generalización del proceso de diagonalización aplicado al caso de las matrices no-diagonalizables. Veamos, en términos generales, en qué consiste.

Definición 5. (Bloque Jordan y matriz Jordan (C. Jordan (1870)))

i) Un *bloque Jordan canónico* de orden n_i para una matriz A de tamaño $n \times n$ es una matriz $n_i \times n_i$, con $n_i \leq n$, de la forma

$$J_{n_i}(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i \end{bmatrix}$$

donde λ_i es un valor propio de A, y n_i es la multiplicidad algebraica de λ_i en el polinomio característico[3].

ii) Una *matriz canónica Jordan* para A es una matriz en la que aparecen bloques Jordan canónicos a lo largo de su diagonal y todas sus otras entradas de la matriz son ceros:

$$\begin{bmatrix} J_{n_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{n_2}(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & J_{n_r}(\lambda_r) \end{bmatrix}$$

Es posible probar (aunque es un tanto difícil en este punto pues requiere de la introducción de muchos conceptos adicionales que no son esenciales para nuestro propósito aquí) que *toda matriz puede reducirse a una matriz canónica de Jordan similar a ella*[4]. Es decir, para cualquier matriz $n \times n$, A, existe una matriz invertible P tal que

$$P^{-1}AP = \begin{bmatrix} J_{n_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{n_2}(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & J_{n_r}(\lambda_r) \end{bmatrix}$$

[3] La multiplicidad algebraica de λ_i en el polinomio característico es el exponente con el que aparece el factor $(\lambda - \lambda_i)$ en el polinomio característico $\det(A - \lambda I_n) = 0$.

[4] Ver Hoffman y Kunze (1971), óp. cit.

donde los λ_i son los valores propios de A y $n_1+n_2+\cdots+n_r=n$. Observemos que si todos los valores propios son reales y distintos, entonces $n_i=1$ para todo i, y $P^{-1}AP$ es, entonces, una matriz diagonal.

También debe advertirse aquí que el cálculo explícito de la matriz P para casos particulares conlleva conceptos y definiciones más allá de los propósitos de este texto.

Finalmente, en el caso concreto de una matriz 2×2 con valores propios *reales*, existe entonces una matriz invertible P tal que $P^{-1}AP=J$, donde J es su matriz Jordan canónica que, así, será de uno de los dos siguientes tipos:

$$\text{i)}\quad \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \qquad \text{ii)}\quad \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$$

donde el caso i) es el de dos valores propios λ_1, λ_2 diferentes dados por $\lambda=\frac{1}{2}(trA\pm\sqrt{(trA)^2-4\det A})$; y el caso ii) es el de un sólo valor propio λ (con multiplicidad 2), con $\lambda=\frac{1}{2}trA$.

Ejemplo 23.
Las siguientes son matrices canónicas de Jordan de alguna matriz A:

$$\text{a)}\quad \begin{bmatrix} 8 & 1 \\ 0 & 8 \end{bmatrix} \qquad\qquad \text{b)}\quad \begin{bmatrix} 7 & 1 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

$$\text{c)}\quad \begin{bmatrix} 4 & 1 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix} \qquad\qquad \text{d)}\quad \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

¿Cuáles son, en cada caso, los valores propios de A y sus respectivas multiplicidades algebraicas?

Ejemplo 24.

a) Para la matriz

$$A=\frac{1}{9}\begin{bmatrix} 20 & -2 & -8 \\ -5 & 14 & 2 \\ 4 & 5 & 20 \end{bmatrix}$$

tenemos que el polinomio característico es $P(\lambda)=(\lambda-2)^3$, y su forma canónica de Jordan es

$$J=\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

b) Para la matriz

$$A = \frac{1}{8}\begin{bmatrix} 15 & 5 & -1 \\ 0 & 2 & 0 \\ 1 & -5 & 17 \end{bmatrix}$$

su forma canónica de Jordan es

$$J = \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

¿Cuál es el polinomio característico y los valores propios de A? ▲

Este teorema, junto con el ejercicio 2 de los Ejercicios 5 adelante, jugará un papel importante en muchas aplicaciones del álgebra lineal a otros problemas matemáticos, en particular, a la teoría de los sistemas dinámicos lineales como entenderemos más adelante (Volumen 3 (Optimización y dinámica); lección 3).

Ejercicios 5

1) Encuentre una matriz canónica de Jordan para

a) $A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$ b) $A = \begin{bmatrix} 1 & 2 & -4 \\ 0 & -1 & 6 \\ 0 & -1 & 4 \end{bmatrix}$

c) $A = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & 2 \\ 1 & 0 & 7 \end{bmatrix}$ d) $A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

2) Si A es de una de las formas canónicas de Jordan

a) $\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ b) $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$

calcule en cada caso A^n para $n = 2, 3, ...$

6. Contexto económico

a. El modelo teórico de Sraffa (1960)

En el "contexto económico" de la lección 3, veíamos cómo el problema de la existencia del valor intrínseco de una mercancía ha estado en el centro de las discusiones de los economistas desde los antiguos griegos hasta, por lo menos, la economía marginalista de finales del siglo XIX, pasando por las discusiones de los clásicos. En 1960, Piero Sraffa [1898-1983] publicó *Production of Commodities by Means of Commodities* (traducido al castellano como *Producción de mercancías por medio de mercancías*) en el cual mostraba cómo sí era posible encontrar, para cualquier economía, una *medida invariable de valor*. Esta mercancía, que se conoce como *mercancía estándar (o patrón)* permite entonces medir *precios relativos que no cambian* excepto cuando la tecnología que la produce, también cambia. Sraffa además demostró que, dados los *salarios reales* y la tecnología de producción, se pueden determinar tanto la tasa de beneficio como los precios relativos. Si los salarios reales están determinados en términos de la mercancía estándar, entonces puede determinarse una *relación lineal* entre el salario real y la tasa de beneficio. Más aún, en el esquema de Sraffa la demanda no entra en el sistema y los precios están determinados por el costo de producción. *Conclusiones como éstas han permitido exploraciones alternas a la teoría walrasiana estándar, en particular en la teoría del valor, en la teoría de la distribución y en la teoría del crecimiento.*

Adelante presentamos una primera aproximación al modelo sraffiano, con énfasis en sus modelos de producción de subsistencia y de producción con excedente, siempre utilizando las técnicas del álgebra lineal[5]. El propósito de esta lección no es, sin embargo, presentar un desarrollo completo de la teoría del valor desde el punto de vista de Sraffa pues esto sobrepasaría con creces los objetivos de este contexto económico. Solamente se desea mostrar cómo pensó el problema regido por una *estructura lineal*, y las implicaciones que conllevaron hacerlo.

i). Supuestos básicos

Las hipótesis básicas consideradas por Sraffa se pueden sintetizar como sigue:

a) El sistema económico se encuentra en condiciones estacionarias y produce cada "año" la misma cantidad de mercancías.

[5] El modelo presentado sigue de cerca el trabajo de Luigi Pasinetti, "El trabajo teórico de Sraffa", en *Lecciones de Teoría de la Producción*, (1984), México: Fondo de Cultura Económica.

b) Los métodos de producción son tales que cada industria produce una sola mercancía mediante el uso de cantidades determinadas de trabajo y de mercancías, las cuales se consumen en su totalidad durante el período correspondiente, por lo que deben ser reemplazadas al final de cada período. De esta manera, al final de cada período, la producción total se divide en la parte que van a reemplazar las mercancías que se utilizaron en el proceso productivo, y la parte que se destina al consumo final.

Los $(n-1)$ métodos de producción posibles se representan en la matriz A de la forma

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1(n-1)} \\ a_{21} & a_{22} & \cdots & a_{2(n-1)} \\ \vdots & \vdots & \vdots & \vdots \\ a_{(n-1)1} & a_{(n-1)2} & \cdots & a_{(n-1)(n-1)} \end{bmatrix} \tag{1}$$

donde los a_{ij} corresponden a los coeficientes interindustriales estándar.

A su vez, el vector fila de coeficientes de trabajo, a_n, será

$$a_n = [a_{n1}, a_{n2}, ..., a_{n(n-1)}] \tag{2}$$

La *técnica del sistema* estará representada entonces por el vector

$$\begin{bmatrix} A \\ a_n \end{bmatrix}$$

Sraffa establece una diferenciación entre mercancías *base* y *no-base*, que depende de las características de los procesos técnicos de producción.

> El criterio consiste en si una mercancía entra (directa o indirectamente) en la producción de *todas* las mercancías. Las que lo hacen serán denominadas productos *básicos*, y las que no lo hacen serán denominadas productos *no básicos*. Supondremos siempre que cualquier sistema contiene al menos un producto básico.

II). El sistema de precios

Sea w el salario unitario y π el tipo de beneficio; dada la técnica del sistema $\begin{bmatrix} A \\ a_n \end{bmatrix}$ y la hipótesis de distribución entre salarios y beneficios, el sistema de precios se define mediante el siguiente sistema lineal de ecuaciones:

$$\begin{aligned}
&(a_{11}p_1 + a_{21}p_2 + \cdots + a_{(n-1)1}p_{n-1})(1+\pi) + a_{n1}w = p_1 \\
&(a_{12}p_1 + a_{22}p_2 + \cdots + a_{(n-1)2}p_{n-1})(1+\pi) + a_{n2}w = p_2 \\
&\vdots \\
&(a_{1(n-1)}p_1 + a_{2(n-1)}p_2 + \cdots + a_{(n-1)(n-1)}p_{n-1})(1+\pi) + a_{n(n-1)}w = p_{n-1}
\end{aligned} \tag{3}$$

donde $p_1, p_2, ..., p_{n-1}$ indican los precios de las mercancías $1, 2, ..., (n-1)$. En forma matricial, el sistema se puede escribir como

$$(1+\pi)AP + wa_n^T = P \tag{3'}$$

donde $P^T = (p_1, p_2, ..., p_{n-1})$.

Resolver este sistema significa encontrar los niveles de precios, el salario y el beneficio para cada uno de los $(n-1)$ métodos de producción posible. El planteamiento se reduce, entonces, a un sistema de $(n-1)$ ecuaciones y $(n+1)$ incógnitas (w, π, y los $(n-1)$ precios). Una forma de resolver esta inconsistencia y encontrar solución al sistema es fijar dos de las incógnitas como, por ejemplo, uno de los precios y el salario, o uno de los precios y el beneficio.

En su trabajo, Sraffa examina las características de las posibles soluciones del sistema (3) tomando como fijo el tipo de beneficio, y para ello estudia todos los valores que éste puede tomar. Al respecto Sraffa explica:

> La dificultad no puede superarse asignando el excedente *antes* de que los precios sean determinados, como se hace con el reemplazamiento de las materias primas, bienes de subsistencia, etcétera. Esto se debe a que el excedente (o beneficio) debe ser distribuido en proporción a los medios de producción (o capital) avanzados en cada industria, y tal proporción entre dos agregados de bienes heterogéneos (en otras palabras, el tipo de beneficio) no puede ser determinada antes de que conozcamos los precios de los bienes. Por otra parte, no podemos diferir la asignación del excedente hasta después de que conozcamos los precios, porque, como veremos, los precios no pueden determinarse antes de conocer el tipo de beneficio. El resultado es que la distribución del excedente debe ser determinada a través del mismo mecanismo y al mismo tiempo que se determinan los precios de las mercancías.

Sraffa considera separadamente dos casos extremos antes de proceder al estudio general:

a) Comienza considerando el caso extremo en que $\pi = 0$. Aquí, el valor de la producción (excedente del sistema) corresponde completamente a los salarios, de acuerdo con la hipótesis c) (el valor del producto se distribuye al final del período en salarios y beneficios). Con uno de los precios y el beneficio fijos, el sistema (3) se convierte en un sistema de $(n-1)$ ecuaciones lineales con n incógnitas (el salario y los $(n-1)$ precios). Para igualar el número de ecuaciones con el número de incógnitas se puede elegir una de las mercancías como numerario (fijando su precio igual a la unidad), quedando el sistema resuelto; es decir, el sistema determina de forma unívoca el salario unitario y los otros $(n-2)$ precios, todos ellos en términos de la mercancía numerario. De igual manera, puede tomarse el salario como numerario ($w = 1$), y el sistema determinaría unívocamente los $(n-1)$ precios, todos ellos en términos del salario.

Con $\pi = 0$, el sistema (3') será entonces

$$(I - A)P = wa_n^T \tag{4}$$

Si $(I - A)$ es una matriz no-singular y multiplicamos ambos miembros de (4) por $(I - A)^{-1}$ se obtiene

$$P = w(I - A)^{-1}a_n^T \tag{5}$$

que, en el caso particular en que $w = 1$, se convierte en

$$P = (I - A)^{-1}a_n^T \tag{6}$$

Dado que a_n es un vector no-negativo, si A satisface las condiciones Hawkins-Simon se puede deducir que $(I - A)^{-1}$ existe y es no-negativa. Por tanto, los precios son todos no-negativos en el sistema (6).

La expresión $(I - A)^{-1}a_n^T$ representa, en términos económicos, las cantidades físicas de las mercancías que han sido utilizadas (directa o indirectamente) en todo el sistema económico para obtener una unidad física de la i-ésima mercancía como mercancía final. De forma similar, el vector v definido como

$$v = (I - A)^{-1}a_n^T \tag{7}$$

que representa los *coeficientes de trabajo verticalmente integrados* (denominadas *cantidades de trabajo "incorporado" en cada unidad física de las mercancías* por los clásicos, y *valores* por Marx) no es más que la representación de la cantidad de trabajo que ha sido necesaria (directa o indirectamente) en todo el sistema económico para obtener una unidad física de la i-ésima mercancía como mercancía final.

De la igualdad (5) tenemos que cuando $\pi = 0$, los precios son proporcionales a las cantidades físicas de "trabajo incorporado" y, en el caso particular en que $w = 1$, los precios son iguales a dichas cantidades físicas de trabajo, siendo los beneficios nulos y, por tanto, todo el producto neto es destinado a los salarios.

b) Examinemos ahora el caso extremo donde el tipo de beneficio alcanza un nivel tal que se anula el salario unitario de forma que todo el valor de la producción corresponda a los beneficios. Denotemos el máximo beneficio (cuando el salario es 0) posible como

$$\pi = \Pi \tag{8}$$

Para obtener el sistema de precios correspondiente a este caso, reemplacemos $w = 0$ en el sistema (3'), de forma que

$$(1 + \Pi)AP = P \tag{9}$$

que es equivalente a escribir

$$((1 + \Pi)A - I)P = 0 \tag{10}$$

Por facilidad llamemos $\chi = \dfrac{1}{1 + \Pi}$, de forma tal que el sistema es ahora

$$(A - \chi I)P = 0 \tag{11}$$

que es un sistema lineal homogéneo. Ya sabemos que la condición necesaria para que este sistema lineal tenga soluciones distintas de cero es que el determinante de la matriz de los coeficientes, es decir, $\det(\chi I - A)$ sea nulo. Encontremos entonces los valores de χ que satisfagan tal condición resolviendo la ecuación característica

$$det\ (A - \chi I) = 0 \tag{12}$$

Las raíces de la ecuación (12) son, sabemos, los autovalores de la matriz A. Hay, en este caso, $(n - 1)$ autovalores con la posibilidad de que algunos estén repetidos. Si A es una matriz no-negativa se puede demostrar (teorema de Frobenius en el Volumen 3 (Optimización y dinámica)) que sólo uno de los $(n - 1)$ autovalores (el máximo autovalor) nos asegura un autovector cuyos componentes son todos no-negativos (para efectos de la interpretación económica y, dado que el autovector son los precios, excluimos los precios negativos del análisis). De esta manera, Π será el tipo de beneficio asociado al máximo autovalor.

Sustituyendo este tipo de beneficio Π en (10) y el máximo autovalor χ^* en (12), el sistema de ecuaciones que se obtiene es lineal y homogéneo con determinante nulo. En el caso en que $w = 0$ se debe satisfacer la siguiente condición:

$$\Pi = \frac{1}{\chi^*} - 1$$

El único tipo de beneficio que nos asegura precios no-negativos (con significado económico) es $\Pi \geq 0$, o bien, $\chi^* \leq 1$. Si esta última condición no se satisface, nos encontraríamos ante un sistema económico que no puede generar beneficios a pesar de que el salario sea cero. Un sistema así no podría, evidentemente, sobrevivir.

Pero ¿qué tipo de teoría del valor está implícita en los precios p^* correspondientes a un salario igual a cero? Una interpretación es considerar el salario unitario determinado por las necesidades de subsistencia, de manera que podemos considerar los bienes que integran el *salario de subsistencia* incluidos en la matriz de coeficientes técnicos, de modo que $w = 0$ significa que el *salario excedente* es igual a cero. En su *Producción de mercancías por medio de mercancías*, Sraffa aclara, con respecto a los salarios, lo siguiente:

> (...) Debemos tener ahora en cuenta el otro aspecto de los salarios, puesto que además del elemento de subsistencia, que siempre está presente en ellos, pueden incluir una participación en la producción excedente. A la vista de este doble carácter de los salarios, sería apropiado, cuando vengamos a considerar la división del excedente entre capitalistas y trabajadores, separar las dos partes componentes del salario y considerar sólo la parte del "excedente" como variable; en tanto que los bienes necesarios para la subsistencia de los trabajadores continuarán apareciendo entre los medios de producción (...)

c) Pasemos ahora a considerar el caso general donde el tipo de beneficio está entre los dos extremos antes expuestos. Sea $\pi = \pi'$ tal que $0 < \pi' < \Pi$; podemos reescribir la ecuación $(3')$ como

$$a_n[I - (1 + \pi')A]^{-1} w = P \tag{13}$$

Para que el sistema tenga solución podemos, nuevamente, tomar el salario igual a 1, de forma que el último sistema es ahora

$$a_n[I - (1 + \pi')A]^{-1} = P \tag{14}$$

Bajo el supuesto de que $\pi' > 0$, $a_n > 0$ y $\pi' < \Pi$, los precios obtenidos son no-negativos (Teorema de Frobenius, Volumen 3 (Optimización y dinámica)). Cuando $\pi' = 0$ se puede comprobar que los precios se hacen proporcionales a las cantidades de trabajo incorporado. Cuando $\pi' > 0$, las cantidades de trabajo indirecto tienen un peso mayor con respecto a las cantidades de trabajo directo. De forma general, tenemos que la estructura de precios depende, tanto de los coeficientes técnicos, como del nivel de beneficio. De acuerdo con la técnica del sistema, existe entonces una estructura de precios correspondiente a cada nivel de beneficio.

III). Los precios y el tipo de beneficio

Aunque no es simple determinar cómo varían los precios cuando cambia el tipo de beneficio, se pueden hacer algunas aproximaciones al respecto. Cuando tomamos como numerario el salario ($w = 1$) se puede concluir, a partir de la ecuación (14), que todos los precios aumentan cuando aumenta π. Sin embargo, en el caso extremo (e hipotético) de mercancías que sólo necesitan trabajo directo y ninguna otra mercancía, los precios permanecen constantes. Tenemos entonces que al aumentar π los precios aumentan o permanecen constantes, pero algunos aumentarán en mayor medida que otros, de manera que, cada precio, relativo a los otros, aumentará o disminuirá.

Si reemplazamos $w = 1$ en la ecuación (3) obtenemos la siguiente relación:

$$\frac{p_j}{p_1} = \frac{a_{nj} + (1+\pi)\sum_{i=1}^{n-1} a_{ij}p_i}{a_{n1} + (1+\pi)\sum_{i=1}^{n-1} a_{i1}p_i} \tag{15}$$

para $j = 2, 3, ..., n-1$. La expresión anterior indica el precio de la j-ésima mercancía en términos de la mercancía 1. La derivada de (15) respecto a π puede ser positiva o negativa; es decir, el precio de la mercancía j respecto al precio de la mercancía 1 puede aumentar o disminuir cuando se incrementa π, dependiendo de que los efectos *de intensidad del capital*[6] y *precio* sean positivos o negativos. El *efecto de intensidad de capital* será positivo cuando las mercancías son producidas con procesos técnicos cuya intensidad en capital sea superior al de la mercancía 1, y será negativo en caso contrario. Por su parte, el *efecto precio* será positivo o negativo dependiendo de las relaciones entre los sectores industriales (depende de todo el sistema económico). A pesar de que estos efectos se pueden reforzar o contrarrestar, se podría argumentar que, en la mayoría de casos, la variación de los precios bajo un determinado tipo de beneficio depende de la intensidad de capital de los procesos productivos.

[6] El efecto de intensidad del capital se representa como $\left[p_1 \sum_{i=1}^{n-1} a_{ij}p_i - p_j \sum_{i=1}^{n-1} a_{i1}p_i\right] + (1+\pi)\left[p_1 \sum_{i=1}^{n-1} a_{ij}\frac{dp_i}{d\pi} - p_j \sum_{i=1}^{n-1} a_{i1}\frac{dp_i}{d\pi}\right]$

Siguiendo a Sraffa,

> La clave del movimiento de precios relativos que sigue a una variación en el salario consiste en la desigualdad de las proporciones en que el trabajo y los medios de producción son empleados en las distintas industrias.
>
> Es claro que si la proporción fuera la misma en todas las industrias no podría seguirse variación alguna de precios por grande que fuera la diversidad de la composición-mercancía de los medios de producción de las diferentes industrias. ...

IV). El salario y el tipo de beneficio

Cuando los precios cambian al variar el tipo de beneficio, la relación *salario-beneficio* depende de dos fenómenos: el primero, es la variación de la distribución de la renta entre salarios y beneficios; y el segundo, la variación de la estructura de precios al variar la distribución de la renta.

Sean α y $(1-\alpha)$ las proporciones de la renta correspondientes a los salarios y beneficios. Bajo el supuesto de que los precios permanecen constantes cuando cambia la distribución de la renta, obtenemos que $\pi = \Pi(1-\alpha)$ (o bien $\pi = \Pi(1-w)$ cuando se toma como numerario el producto neto por trabajador); es decir, una relación lineal entre el beneficio y el salario.

A medida que el salario se reduce gradualmente de 1 a 0, el tipo de beneficio aumenta en proporción directa a la deducción total hecha del salario. La relación puede ser representada por una línea recta, tal como aparece en la figura 6.

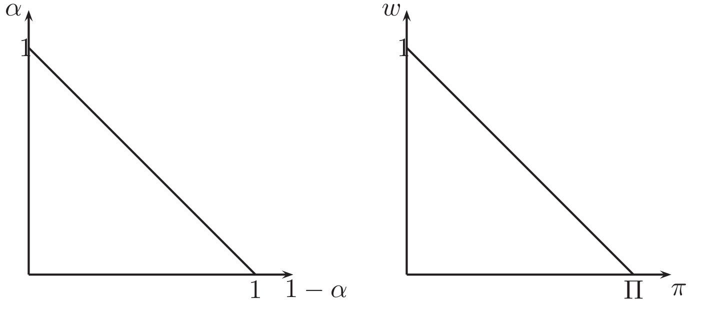


Figura 6

Si tomamos el caso más general, en el que los precios pueden expresarse en términos de una mercancía cualquiera como numerario, el salario (en términos de cualquier mercancía) es siempre una función monótona decreciente del tipo de beneficio. Se trata de una función polinómica que adopta una forma distinta según la mercancía escogida como numerario, lo que hace difícil concluir con respecto a la distribución de la renta entre salarios y beneficios.

Una manera de resolver esta dificultad es reducir los precios a cantidades de trabajo fechada. La matriz $[I-(1+\pi)A]^{-1}$ se puede descomponer en una serie de potencias de A, con la condición de que $\pi < \Pi$, donde $\frac{1}{\Pi}+1$ es el máximo autovalor de A (Volumen 3 (Optimización y dinámica)). Podemos entonces reescribir el vector de precios como

$$p = a_n[(I + (1+\pi)A + (1+\pi)^2A^2 + (1+\pi)^3A^3 + \cdots)] \tag{16}$$

Cuando $\pi = 0$ y $w = 1$ la expresión anterior se reduce a

$$p = a_n + a_nA + a_nA^2 + a_nA^3 + \cdots \tag{17}$$

que es una serie de vectores que representan los requisitos de trabajo en las sucesivas fases del proceso productivo, donde el primer sumando de la expresión corresponde a las cantidades de trabajo directo y la suma de los demás a las cantidades de trabajo indirecto. Podemos concluir de (16) que todos los precios aumentan al aumentar el beneficio. En el límite, la suma crece indefinidamente y los precios, en términos del salario, también crecen infinitamente. Para los tipos de beneficio en el intervalo $0 \leq \pi < \Pi$ la suma se aproxima a un valor fijo, y proporciona la reducción de todos los precios a cantidades de trabajo fechadas oportunamente ponderadas mediante el correspondiente $(1+\pi)^t$ ($t =$ etapa de producción) de capitalización compuesta.

Si $w \neq 1$ y el numerario es cualquier mercancía, podemos escribir la ecuación (16) en la forma

$$p = a_nw + (1+\pi)a_nAw + (1+\pi)^2a_nA^2w + \cdots \tag{16'}$$

Dado que w y π pueden variar en dirección opuesta, algunos sumandos aumentan y otros disminuyen, mientras que cada precio aumentará o disminuirá como efecto de la suma de estas variaciones.

Bajo la interpretación de la ecuación ($16'$) como los distintos “estratos” (a medida que retrocedemos en el tiempo) de beneficio y salario que constituyen el precio de una mercancía, la causa de los cambios en los precios al variar la distribución de la renta se encuentra en las distintas proporciones entre trabajo y medios de producción requeridos por cada mercancía en los distintos “estratos” que integran el precio.

v). El sistema de cantidades

Veamos ahora cuál es el sistema de ecuaciones correspondiente a las cantidades de mercancías que se van a producir. De forma similar al sistema de precios, podemos escribir el sistema de cantidades como

$$\begin{aligned}
&a_{11}Q_1 + a_{12}Q_2 + \cdots + a_{1(n-1)}Q_{n-1}Y_1 = Q_1\\
&a_{21}Q_1 + a_{22}Q_2 + \cdots + a_{2(n-1)}Q_{n-1}Y_2 = Q_2\\
&\vdots\\
&a_{(n-1)1}Q_1 + a_{(n-1)2}Q_2 + \cdots + a_{(n-1)(n-1)}Q_{n-1}Y_{n-1} = Q_{n-1}
\end{aligned}\tag{18}$$

En este caso, $Y_1, Y_2, ..., Y_{n-1}$, las mercancías finales, son datos conocidos.

Expresemos las cantidades de mercancías destinadas al sector final (consumo final) en términos relativos, esto es, en términos del tipo de excedente físico que notaremos como R_i para $i = 1, 2, ..., n-1$, y que definimos como sigue:

$$R_1 = \frac{Y_1}{Q_1 - Y_1}, R_2 = \frac{Y_2}{Q_2 - Y_2}, \quad ..., \quad R_{n-1} = \frac{Y_{n-1}}{Q_{n-1} - Y_{n-1}}$$

El sistema de cantidades puede entonces reescribirse como

$$\begin{aligned}
&(a_{11}Q_1 + a_{12}Q_2 + \cdots + a_{1(n-1)}Q_{n-1})(1 + R_1) = Q_1\\
&(a_{21}Q_1 + a_{22}Q_2 + \cdots + a_{2(n-1)}Q_{n-1})(1 + R_2) = Q_2\\
&\vdots\\
&(a_{(n-1)1}Q_1 + a_{(n-1)2}Q_2 + \cdots + a_{(n-1)(n-1)}Q_{n-1})(1 + R_{n-1}) = Q_{n-1}
\end{aligned}\tag{19}$$

Ahora: $R_1, R_2, ..., R_{n-1}$ son los datos conocidos. El problema es encontrar las $Q_1, Q_2, ..., Q_{n-1}$ que deben producirse para obtener las cantidades finales.

Podemos reescribir el sistema anterior como:

$$\begin{bmatrix}
(a_{11} - \frac{1}{1+R_1}) & a_{12} & \cdots & a_{1(n-1)}\\
a_{21} & (a_{22} - \frac{1}{1+R_2}) & \cdots & a_{2(n-1)}\\
\vdots & \vdots & \vdots & \vdots\\
a_{(n-1)1} & & \cdots & (a_{(n-1)(n-1)} - \frac{1}{1+R_{n-1}})
\end{bmatrix}
\begin{bmatrix} Q_1\\ Q_2\\ \vdots\\ Q_{n-1}\end{bmatrix}
= \begin{bmatrix} 0\\ 0\\ \vdots\\ 0\end{bmatrix}\tag{20}$$

Este sistema lineal es homogéneo y, por tanto, la condición necesaria para que hayan soluciones distintas de cero es que el determinante de la matriz de los coeficientes sea nulo. Por tanto, al menos una de las R_i deberá ser tal, que haga que se cumpla con la condición requerida del determinante. Tal sistema dará soluciones para las cantidades físicas $Q_1, Q_2, ..., Q_{n-1}$ en función de una constante que las multiplica, por lo que quedará determinada la *estructura de producción* del sistema, pero indeterminada la escala de producción.

Así, dado que todas las R_i excepto una son consideradas como datos, podemos introducir la hipótesis de que todas sean iguales entre sí, esto es

$$R_1 = R_2 = R_3 = \cdots = R_{n-1} = R,$$

donde R es el tipo de excedente uniforme de todo el sistema. Si sustituimos R en (20), y definimos $\dfrac{1}{1+R} = \delta$, obtenemos un sistema de ecuaciones, que podemos escribir en forma matricial como:

$$AQ = \delta Q \tag{21}$$

ó como

$$(A - \delta I)Q = 0$$

La condición necesaria para obtener soluciones diferentes de cero es que el determinante $|\delta I - A| = 0$ que, en general, tendrá $(n-1)$ raíces, aunque es posible que algunas de ellas se repitan. Tales raíces son los autovalores de la matriz A, que es una matriz no-negativa. Tenemos entonces que los valores de δ que satisfacen la igualdad anterior coinciden con los valores de χ que habíamos encontrado antes, y por lo tanto,

$$R = \Pi$$

es decir, *el tipo de excedente uniforme del sistema iguala la tasa de beneficio máximo.* Este importante resultado se obtiene de las propiedades matemáticas de las ecuaciones que representan el sistema productivo. Si la igualdad entre R y Π no se cumpliera, significaría que el sistema económico considerado sería tan obsoleto que originaría un tipo de excedente negativo; esto es, un sistema tal que necesitaría, como medios de producción, cantidades de mercancías superiores a las que estaría en disposición de producir.

Determinado el excedente R, éste se introduce en el sistema (21) y así se obtienen las Q_i que dependen de una constante ((21) representa un sistema lineal y homogéneo con determinante nulo). Para calcular la escala de producción es necesario introducir como supuesto adicional que la cantidad de trabajo existente esté dado, de forma que $a_n Q = \overline{Q_n}$. Resolviendo esta igualdad junto con $(1+R)AQ = Q$ obtenemos las Q_i^*. El teorema de Frobenius (volumen 3 (Optimización y dinámica)) asegura que las soluciones son no-negativas (positivas o cero). Podemos, por tanto, eliminar de la matriz de coeficientes las filas y columnas correspondientes. Si las soluciones son positivas obtenemos que

$$(Q_k^* - Y_k^*) = \frac{R}{(1+R)}[Q_1^*, Q_2^*, ..., Q_k^*]$$

De esta manera, *las cantidades positivas correspondientes a las soluciones del sistema son tales que las proporciones en las que las distintas mercancías son producidas, son iguales a las proporciones en que estas aparecen como medios de producción en el sistema económico y, además, iguales a las proporciones en que dichas mercancías son destinadas a los sectores finales (de consumo).*

Nota 4.
Sraffa realizó su trabajo con paciencia y dedicación ejemplar. Es así como su libro *Producción de mercancías por medio de mercancías*, aunque tiene origen en un manuscrito de 1928, sólo fue publicado hasta 1960. De forma similar, su edición de las *Obras completas* de David Ricardo, cuya publicación se anunció en 1933, sólo se terminó hasta 1951. Su obra fue un intento por perfeccionar la teoría clásica del valor desarrollada originalmente por David Ricardo y, en este sentido, intentó demostrar las fallas de la teoría neoclásica del valor y desarrollar un análisis alternativo. El trabajo de Sraffa ha servido de inspiración a algunos economistas poskeynesianos para justificar el abandono del análisis neoclásico. Otros, sin embargo, ven su trabajo compatible con el análisis neoclásico, como, por ejemplo, con los modelos de equilibrio general walrasiano (Hahn (1982))[7].

[7] Hahn, Frank (1982), The Neo-Ricardians, *Cambridge Journal of Economics*, vol. 6, 353-374.

Ejercicios complementarios

1) Calcule los valores propios de la transformación $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por

$$T(\,x_1, x_2\,) = A\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \text{donde } A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

es decir, $T(\,x_1, x_2\,) = (\,-x_2,\, x_1\,)$. ¿Qué representa geométricamente esta transformación lineal?

2) Encuentre una matriz invertible P tal que $P^{-1}AP$ sea diagonalizable si

a) $A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$ b) $A = \begin{bmatrix} 3 & -8 \\ 0 & -2 \end{bmatrix}$

3) Si $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ y $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, encuentre $P^{-1}AP$. Interprete geométricamente.

4) Diagonalice, si es posible, la matriz

$$A = \begin{bmatrix} 1 & 3 & 0 & 0 \\ -3 & -1 & 1 & 0 \\ -3 & -1 & 2 & 1 \\ 1 & 1 & -1 & 0 \end{bmatrix}$$

5) (Hoffman y Kunze (1971)) Sea T un operador lineal sobre $\mathbb{R}^4$ que, representado en la base canónica, aparece como la matriz

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \end{bmatrix}$$

¿Bajo qué condiciones de a, b, c (si existen) es T diagonalizable?

6) Muestre que si

$$A = \begin{bmatrix} 7.3 & 0.2 & -3.7 \\ -11.5 & 1.0 & 5.5 \\ 17.7 & 1.8 & -9.3 \end{bmatrix}$$

entonces A es diagonalizable mediante la matriz P, donde

$$P = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix}$$

7) Pruebe que la matriz P que diagonaliza la matriz

$$A = \begin{bmatrix} -17 & 18 & -6 \\ -18 & 19 & -6 \\ -9 & 9 & -2 \end{bmatrix}$$

es

$$P = \begin{bmatrix} 2 & 1 & -1 \\ 2 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

¿Cuál es $P^{-1}AP$?

8) Para cada una de las siguientes matrices, encuentre la forma diagonal y la matriz P que la diagonaliza:

a) $\begin{bmatrix} 3 & 4 \\ -4 & 3 \end{bmatrix}$ b) $\begin{bmatrix} \operatorname{sen}\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$, θ dado

c) $\begin{bmatrix} 20 & 18 \\ -27 & -25 \end{bmatrix}$ d) $\begin{bmatrix} 1 & 0 & 0 & -5 \\ 0 & -1 & 0 & 6 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

9) Pruebe que las siguientes matrices no son diagonalizables:

a) $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 2 & -4 \\ 0 & -1 & 6 \\ 0 & -1 & 4 \end{bmatrix}$

10) Pruebe que la forma cuadrática

$$Q(\,x_1, x_2\,) = 17x_1^2 - 30x_1x_2 + 17x_2^2 = 128$$

representa la elipse $2y_1^2 + 32y_2^2 = 128$, transformando los ejes principales.

**11) Demuestre que una ecuación de segundo grado $Ax^2 + By^2 + Cxy + Dx + Ey + F = 0$ está dentro de una (y sólo una) de las 9 clases canónicas:

a) La elipse $\left(\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1\right)$

b) La elipse imaginaria [8] $\left(\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = -1\right)$

---

[8] ¿Podría el lector dar aquí una justificación del término "imaginario"? (Volumen 0 (Fundamentos)), lección 3.

c) El punto $\left(\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=0\right)$

d) La hipérbola $\left(\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=1\right)$

e) Un par de líneas que se intersecan $\left(\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=0\right)$

f) La parábola $(\,y^2=2ax,\;\; a>0)$

g) $x^2-a^2=0$ (par de líneas paralelas)

h) $x^2+a^2=0$ (par de líneas paralelas imaginarias)

i) $x^2=0$ (par de líneas rectas coincidentes)

[Indicación: este ejercicio es de un nivel superior aún para el estudiante aventajado.]

12) Asuma que una forma cuadrática de la forma

$$A_1x^2+A_2y^2+A_3z^2+2B_1yz+2B_2xz+2B_3xy+2C_1x+2C_2y+2C_3z+D=0$$

está dentro de una (y sólo una) de las 17 clases afines canónicas de formas cuadráticas:

a) Elipsoide: $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=1$

b) Elipsoide imaginario: $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=-1$

c) Hiperboloide de una hoja: $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=1$

d) Hiperboloide de dos hojas: $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=1$

e) Cono de segundo orden: $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=0$

f) Cono imaginario de segundo orden: $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=0$

g) Paraboloide elíptico: $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}-\dfrac{z}{c}=0$

h) Paraboloide hiperbólico: $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}-\dfrac{z}{c}=0$

i) Cilindro elíptico: $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$

j) Cilindro elíptico imaginario: $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+1=0$

k) Cilindro hiperbólico: $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$

l) Cilindro parabólico: $y^2 - 2ax = 0$

m) Par de planos que se intersecan: $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 0$

n) Par de planos imaginarios que se intersecan: $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 0$

o) Par de planos paralelos: $x^2 - a^2 = 0$

p) Par de planos imaginarios: $x^2 + a^2 = 0$

q) Par de planos coincidentes: $x^2 = 0$

Determine a cuál de los 17 tipos anteriores pertenecen las cuadráticas siguientes:

a) $2xy + 2yz + 2xz = 2$

b) $2x^2 + z^2 + 2xy + 3xz + yz = 2$

c) $x^2 - yz - 2x - 4y - 3z = 5$

13) a) Muestre que si $Q = XAX^T$ es definida positiva, entonces la forma $Q^* = XA^{-1}X^T$ también es definida positiva. Ilustre con un ejemplo.

b) Muestre que si XAX^T y XBX^T son dos formas $n \times n$ simétricas definidas positivas, entonces también la forma $X(B^{-1} - A^{-1})X^T$ es definida positiva. Ilustre con un ejemplo.

14) Muestre que si A es una matriz $n \times m$, $n \geq m$, $\rho(A) = m$, entonces la matriz $P = A(A^TA)^{-1}A^T$ satisface:

a) $P = P^T = P^2$ (es decir, P es simétrica e *idempotente*).

b) $\rho(P) = m$ (rango de P es igual a m).

c) Las raíces características de P consisten de m unos y $n - m$ ceros.

15) Muestre que una matriz Jordan para

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{es} \quad J = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

donde $J = P^{-1}AP$ para la matriz inversible

$$P = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

16) Encuentre una matriz Jordan para

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

17) Para tener una comprensión más profunda de las transformaciones lineales (y de su diagonalización) es necesario entender el significado geométrico de los valores propios *complejos* de una transformación, cuando estos surgen. Para intentar alcanzar esto, recordemos primero (Volumen 0 (Fundamentos)) que el conjunto de los números complejos está definido por

$$\mathbb{C} = \{a + ib/a, b \in \mathbb{R}, i^2 = -1\}$$

y que $\mathbb{C}$ tiene la misma estructura algebraica de $\mathbb{R}$ con respecto a la suma y el producto.

a) Pruebe que las transformaciones lineales $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definidas por

$$T(x, y) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} (\equiv AX)$$

para $b > 0$, describen una rotación de un ángulo de θ radianes en sentido contrario a las manecillas del reloj, seguido por una contracción r de ejes, donde

$$r = \sqrt{a^2 + b^2} \quad \text{y} \quad \cos\theta = \frac{a}{r}$$

[Indicación: utilice las coordenadas polares $a = r\cos\theta$,
$b = r \operatorname{sen}\theta$ y muestre que

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix} \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}]$$

b) Muestre que las raíces del polinomio característico de A *no son números reales*: son los *números complejos* $\lambda_1 = a + ib$, y su conjugado $\lambda_2 = a - ib$ donde i es tal que $i^2 = -1$.

En general, sabemos (volumen 0 (Fundamentos)) que si $\lambda = a+ib \in \mathbb{C}$ es una raíz de un polinomio con coeficientes reales, entonces su conjugado $\overline{\lambda} = a - ib$ también es una raíz de él. Si aplicamos esto al polinomio característico de una transformación lineal $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ encontraremos que si T presenta valores propios complejos entonces implica, dentro de sus “movimientos rígidos”, rotaciones y contracciones de ejes.

18) Sea A una matriz $n \times n$ con entradas reales. Si λ es un valor propio *complejo* de A entonces un correspondiente vector propio X tendrá entradas complejas. Pruebe que, por consiguiente, un vector propio asociado a $\overline{\lambda}$ será el vector $\overline{X}$ conformado en sus entradas por los conjugados de las entradas de X. Es decir, pruebe que si $AX = \lambda X$ entonces $A\overline{X} = \overline{\lambda X}$.

    [Indicación: tome conjugados a ambos lados de la igualdad $AX = \lambda X$].

19) (*Demostración de la parte a) del teorema 6*) Asuma que $AX = \lambda X$ y que (por el problema 18) $A\overline{X} = \overline{\lambda X}$ para algún $\lambda \in \mathbb{C}$.

    a) Pruebe que $\overline{X}^T AX = \lambda \overline{X}^T X$ y $X^T A\overline{X} = \lambda X^T \overline{X}$.
    b) Pruebe que $\overline{X}^T AX = (X^T A\overline{X})^T$ aplicando el hecho de que A es simétrica.
    c) Concluya que $\lambda \overline{X}^T X = \overline{\lambda X}^T X$ y que, por tanto, $\lambda = \overline{\lambda}$, y así λ es un número real. ■

    (Nota: la parte b) del teorema 6 es un resultado profundo que requiere un mayor desarrollo temático que está fuera del alcance de este texto)[9].

20) Suponga un sistema económico sraffiano en el que los coeficientes técnicos están representados en la siguiente matriz:

$$A = \begin{bmatrix} 0.20 & 1.28 & 0.25 \\ 0.01 & 0.14 & 0.02 \\ 0.01 & 0.14 & 0.12 \end{bmatrix}$$

    Muestre que el autovalor máximo es $\lambda^* = 0.337$. Calcule el tipo máximo de beneficio y las cantidades relativas del sistema considerado.

$*$22) ¿Qué diferencias sustanciales encuentra usted entre los modelos

    a) Walras-Cassel y Sraffa?
    b) Leontief y Sraffa?
    c) von Neumann y Sraffa?

[9] Ver Hoffman y Kunze (1971), óp. cit.

Lección 8

Conjuntos convexos

Introducción

Hasta donde se sabe, las nociones de convexidad y, por tanto, de concavidad, se presentaban de manera corriente en los intentos de los antiguos griegos por resolver los famosos problemas de la geometría euclidiana como la cuadratura del círculo (construir, sólo con regla y compás, un cuadrado cuya área fuera igual a la de un círculo dado) o la trisección de un ángulo (dividir un ángulo en tres partes iguales sólo con regla y compás). Pero ha sido más de dos mil años después, a partir de los desarrollos de la teoría de la optimización en la década de 1930, que la conveniencia de estos "conjuntos lineales" ha venido a jugar un papel fundamental en el desarrollo de nuevas matemáticas. Hoy su pertinencia es indiscutible, pues los conjuntos convexos surgen en numerosas situaciones: en optimización clásica, por ejemplo, es corriente que el conjunto de puntos posibles de solución forme un conjunto convexo; pero también en estadística los "conjuntos de riesgo" son conjuntos convexos, y también lo es el conjunto de valores esperados de una "variable aleatoria". Estudiemos entonces un poco sobre lo que es convexidad y algunos casos específicos en los que esta noción surge naturalmente, además de su conexión con el tipo de "pensamiento lineal" que aquí nos ha interesado.

1. Noción de conjunto convexo

Intuitivamente, un conjunto convexo es aquel que tiene la propiedad de que siempre que escojamos dos puntos del conjunto, entonces todos los puntos sobre el segmento de recta que los une también estarán en el conjunto. Y este es el concepto central de esta lección.

Definición 1. (Conjunto convexo)
Diremos que un conjunto no vacío $C \subseteq \mathbb{R}^n$ es un *conjunto convexo* si, y sólo

si, para todo $x,\ y \in C$, $\lambda \in [\,0, 1\,]$,

$$\lambda x + (\,1 - \lambda\,)y \in C$$

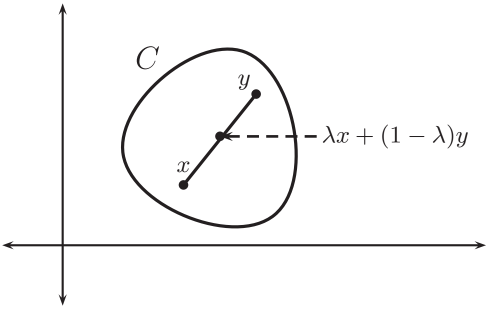


Figura 1. Conjunto convexo

Ejemplo 1.
El círculo de radio r, $C = \{x \in \mathbb{R}^n /\ \| \ x \ \| \leq r\}$ (figura 2), es un conjunto convexo. En efecto, para $x, y \in C$, y $\lambda \in [0, 1]$, se tiene que

$$\begin{aligned} \| \ \lambda x + (1-\lambda)y \ \| &\leq \lambda \ \| \ x \ \| + (1-\lambda) \ \| \ y \ \| \\ &\leq \lambda r + (1-\lambda)r = r \end{aligned}$$

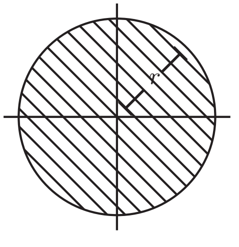


Figura 2. El círculo de radio r

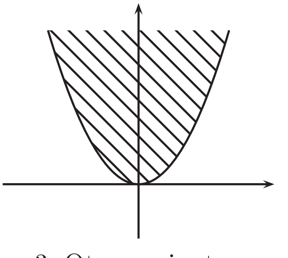


Figura 3. Otro conjunto convexo

Ejemplo 2. (Otro conjunto convexo)
Probemos que el conjunto $C = \{(x, y) \in \mathbb{R}^2 / y \geq x^2\}$ es convexo (figura 3).

Solución

Sean $(x_1, y_1), (x_2, y_2) \in C$, y $\lambda \in [0, 1]$. Entonces

$$\lambda(x_1, y_1) + (1-\lambda)(x_2, y_2) = (\lambda x_1 + (1-\lambda)x_2,\ \lambda y_1 + (1-\lambda)y_2) \in C$$

puesto que, claramente, $(\lambda x_1 + (1-\lambda)x_2)^2 = \lambda^2 x_1^2 + (1-\lambda)^2 x_2^2 + 2\lambda(1 - \lambda)x_1 x_2 \leq \lambda^2 y_1 + (1-\lambda)^2 y_2 + 2\lambda(1-\lambda)\sqrt{y_1}\sqrt{y_2} = (\lambda\sqrt{y_1} + (1-\lambda)\sqrt{y_2})^2 \leq$

$\lambda y_1 + (1-\lambda)y_2$. Que esta última desigualdad es cierta, lo vemos del hecho de que $(\lambda\sqrt{y_1} + (1-\lambda)\sqrt{y_2})^2 \leq \lambda y_1 + (1-\lambda)y_2$ si, y sólo si, $\lambda^2 y_1 + (1-\lambda)^2 y_2 + 2\lambda(1-\lambda)\sqrt{y_1}\sqrt{y_2} \leq \lambda y_1 + (1-\lambda)y_2$; de la cual, reordenando, obtenemos que $\lambda(\lambda-1)y_1 + \lambda(\lambda-1)y_2 + 2\lambda(1-\lambda)\sqrt{y_1}\sqrt{y_2} \geq 0$. Si $\lambda = 0, 1$ hemos finalizado. En otro caso, llegamos a que $y_1 + y_2 - 2\sqrt{y_1}\sqrt{y_2} \geq 0$ o, equivalentemente, $\left(\sqrt{y_1} - \sqrt{y_2}\right)^2 \geq 0$, y esta última desigualdad es siempre cierta.

Nota 1.
¿Podría el lector decir por qué la prueba del ejemplo 1 fue analíticamente más simple que la prueba del ejemplo 2?

Ejemplo 3. (Las roscas no son conjuntos convexos)
El conjunto resaltado en la figura 4 *no* es un conjunto convexo pues el segmento de recta que une los puntos A y B en la figura no está totalmente contenido en el conjunto.

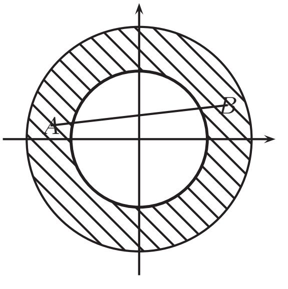


Figura 4. Conjunto no-convexo

Las siguientes son las propiedades centrales de los conjuntos convexos:

Teorema 1. ***(Propiedades básicas de los conjuntos convexos)***
Sean C_1 y $C_2 \subseteq \mathbb{R}^n$ conjuntos convexos, α, $\beta \in \mathbb{R}$; entonces

a) *$C_1 \cap C_2$ es un conjunto convexo; más aún, $\bigcap_{i\in X} C_i$ es un conjunto convexo, para toda familia $\{C_i\}_{i\in X}$ de conjuntos convexos.*

b) *$\alpha C_1 + \beta C_2 = \{\, \alpha c_1 + \beta c_2 \,/\, c_1 \in C_1,\ c_2 \in C_2 \,\}$ es un conjunto convexo.*

c) *$C_1 \times C_2 = \{(c_1, c_2)/c_1 \in C_1, c_2 \in C_2\}$ (producto cartesiano) es un conjunto convexo.*

d) *Si $T : R^n \to R^m$ es una transformación lineal, entonces $T(C_1)$ es un conjunto convexo en R^m.*

Demostración

a) Sean $x,\ y \in C_1 \cap C_2$. Entonces $x,\ y \in C_1,\ \ x,\ y \in C_2$. Luego, para $\lambda \in [\,0,1\,]$, puesto que C_1 y C_2 son convexos, se tendrá que $\lambda x + (\,1-\lambda\,)y \in C_1$, y también, $\lambda x + (\,1-\lambda\,)y \in C_2$. Luego $\lambda x + (\,1-\lambda\,)y \in C_1 \cap C_2$ (figura 5a).

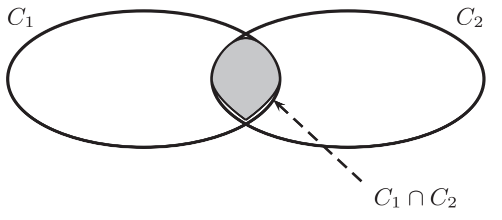


Figura 5a. Intersección de conjuntos convexos

La prueba es similar para una familia de conjuntos convexos.

b) Sean $x,\ y \in \alpha C_1 + \beta C_2$; entonces $x = \alpha c_{11} + \beta c_{21},\ \ y = \alpha c_{12} + \beta c_{22}$ para $c_{11},\ c_{12} \in C_1$ y $c_{21},\ c_{22} \in C_2$. Luego, para $\lambda \in [\,0,1\,]$,

$$\begin{aligned}\lambda x + (\,1-\lambda\,)y &= \lambda(\,\alpha c_{11} + \beta c_{21}\,) + (\,1-\lambda\,)(\,\alpha c_{12} + \beta c_{22}\,)\\ &= \alpha(\,\lambda c_{11} + (\,1-\lambda\,)c_{12}\,) + \beta(\,\lambda c_{21} + (\,1-\lambda\,)c_{22}\,)\\ &\in \alpha C_1 + \beta C_2\end{aligned}$$

pues $\lambda c_{11} + (\,1-\lambda\,)c_{12} \in C_1$ y $\lambda c_{21} + (\,1-\lambda\,)c_{22} \in C_2$.

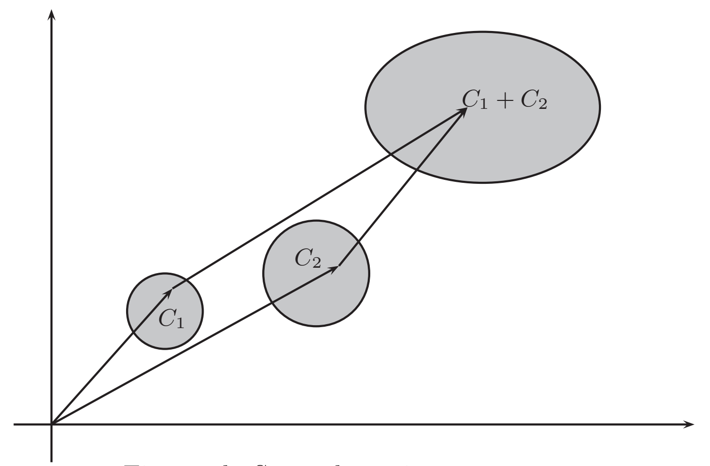


Figura 5b. Suma de conjuntos convexos

d) Si $x, y \in C_1, \lambda \in [0,1]$ entonces $\lambda T(x) + (1-\lambda)T(y) = T(\lambda x + (1-\lambda)y) \in T(C_1)$ pues C_1 es convexo, dado que $\lambda x + (1-\lambda)y \in C_1$.

El literal c) se deja como ejercicio para el lector. ■

Ejemplo 4.
Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación lineal definida por $T(x, y) = (2x, 3y)$. Describamos la imagen bajo T del conjunto convexo $C = \{(x, y) \in \mathbb{R}^2 / \|(x, y)\| \leq 1\}$ (círculo de radio 1).

Solución

Sea $(x, y) \in C$; entonces, si $T(x, y) = (u, v)$ donde $u = 2x$, $v = 3y$, estas u, v satisfacen la desigualdad $\left(\frac{u}{2}\right)^2 + \left(\frac{v}{3}\right)^2 \leq 1$, que es la parte interior de la elipse con semieje menor 2 y semieje mayor 3 (Volumen 0: Fundamentos). Así, una elipse no es más que un círculo "transformado linealmente" (figura 6). ▲

Además de los mismos espacios $\mathbb{R}^n$, los más típicos ejemplos de conjuntos convexos son los *hiperplanos*. Veamos su definición.

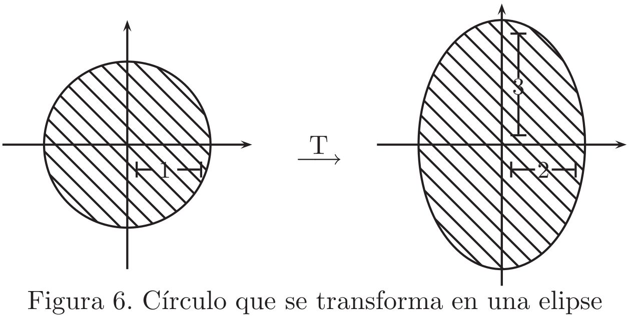


Figura 6. Círculo que se transforma en una elipse

Definición 2. (Hiperplano (afín))
Un conjunto no vacío $H \subseteq \mathbb{R}^n$ es un *hiperplano (afín)* de $\mathbb{R}^n$ si, y sólo si, existe $x_0 \in \mathbb{R}^n$ tal que $H - \{ x_0 \}$ es un subespacio vectorial propio de $\mathbb{R}^n$ (figura 7). Así, es fácil probar que un hiperplano afín H tiene la forma

$$H = \{x \in \mathbb{R}^n / \, p \cdot x = c\}$$

para ciertos $p \in \mathbb{R}^n, \; c \in \mathbb{R}$ fijos (¿Cómo depende c de p y x_0?). Claramente, todo hiperplano afín de $\mathbb{R}^n$ es un conjunto convexo, pues si $x_1, x_2 \in H$ entonces también $\lambda x_1 + (1 - \lambda)x_2 \in H$ con $\lambda \in [0, 1]$; en efecto: $p(\lambda x_1 + (1 - \lambda)x_2) = \lambda p x_1 + (1 - \lambda)p x_2 = \lambda c + (1 - \lambda)c = c$.

Ejemplos euclidianos típicos de hiperplanos son las rectas en el plano $\mathbb{R}^2$ (sin que necesariamente pasen por el origen) y todos los planos en $\mathbb{R}^3$ (sin que necesariamente pasen por el origen).

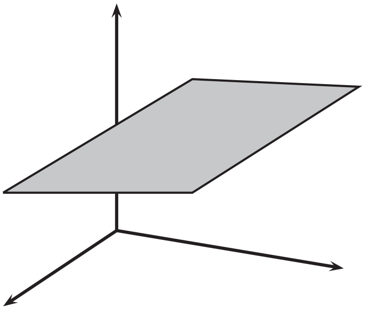


Figura 7. Hiperplano (afín) en $\mathbb{R}^3$

Ahora: dado un hiperplano H, el conjunto H^+ formado por todos los $x \in \mathbb{R}^n$ tales que $p \cdot x \geq c$ es también un conjunto convexo conocido como *semiespacio cerrado superior*; de manera similar, H^-, conocido como *semiespacio cerrado inferior*, es el conjunto de todos los $x \in \mathbb{R}^n$ tales que $p \cdot x \leq c$. La demostración de que H^+ y H^- son convexos es simple y se deja como ejercicio para el lector.

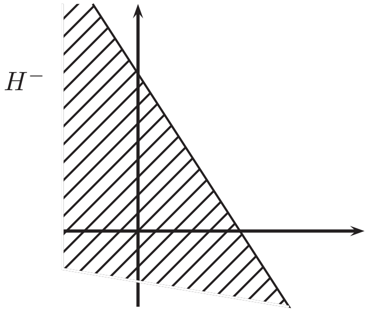


Figura 8a. Semiespacio cerrado inferior en $\mathbb{R}^2$

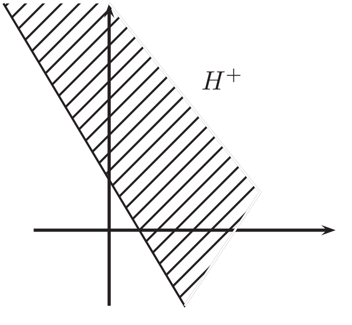


Figura 8b. Semiespacio cerrado superior en $\mathbb{R}^2$

Otras clases importantes de conjuntos convexos son las siguientes:

Definición 3. (Politopos y poliedros)

a) Un *politopo* en $\mathbb{R}^n$ es la intersección de un número *finito* de semiespacios cerrados.

b) Un *poliedro* (o conjunto poliedral) en $\mathbb{R}^n$ es un politopo *acotado*.

De acuerdo con el teorema 1, todo politopo o poliedro es un conjunto convexo.

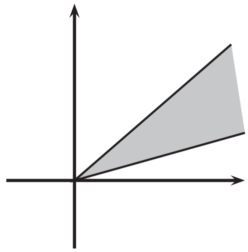


Figura 9a. Politopo en $\mathbb{R}^2$

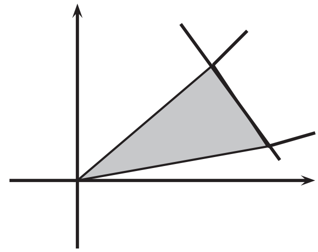


Figura 9b. Poliedro en $\mathbb{R}^2$

Finalmente presentamos la última clase de conjuntos que acostumbra asociarse con la noción de convexidad: *los conos.*

Definición 4. (Conos)
Diremos que $C \subseteq \mathbb{R}^n$ es un *cono (con vértice $x_0 \in \mathbb{R}^n$)*, si contiene todos los puntos de la forma $z = x_0 + t(y - x_0)$ con $t \geq 0$, siempre que contenga el punto y. Si C es un conjunto convexo entonces lo llamaremos un *cono convexo* (figura 10).

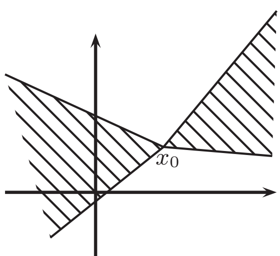


Figura 10a. Cono (no-convexo) con vértice en x_0, en $\mathbb{R}^2$

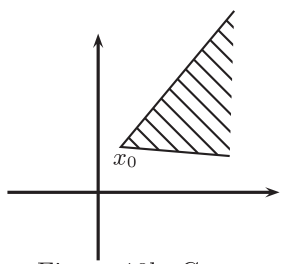


Figura 10b. Cono convexo con vértice en x_0, en $\mathbb{R}^2$

Ejemplo 5.

i) Si $p \in \mathbb{R}^n$, el conjunto $C = \{x \in \mathbb{R}^n / px \leq 0\}$ es un cono convexo.

ii) El conjunto $\{(x, y) \in \mathbb{R}^2_+ / x \leq y\} \bigcap \{(x, y) \in \mathbb{R}^2_+ / x \geq -5y\}$ es un cono con vértice en el origen. ▲

Una operación fundamental a la noción de convexidad, dado su "carácter lineal", es tomar "combinaciones convexas" de los puntos de un conjunto no-convexo dado, para transformarlo en uno convexo. Así, es posible generar nuevos conjuntos convexos a partir de otros que, originalmente, no lo son.

Definición 5. (Combinación convexa)
Una *combinación convexa* de $x_0, x_1, \ldots, x_p \in \mathbb{R}^n$ es un vector $x \in \mathbb{R}^n$ que puede expresarse como

$$x = \lambda_0 x_0 + \lambda_1 x_1 + \cdots + \lambda_p x_p$$

con $\lambda_0 + \lambda_1 + \cdots + \lambda_p = 1;\ \lambda_0, \lambda_1, \ldots, \lambda_p \geq 0.$

Definición 6. (Envolvente convexa)
Si $C \subseteq \mathbb{R}^n$ es cualquier subconjunto no vacío de $\mathbb{R}^n$, la *envolvente (o envoltura) convexa* de C es el conjunto de todas las posibles combinaciones convexas entre puntos de C, y lo notaremos por $\mathrm{Conv}(C)$.

Teorema 2. ***(¿Cómo "convexificar" un conjunto cualquiera?)***
La envolvente convexa de cualquier conjunto es un conjunto convexo. Más aún, es el mínimo conjunto convexo que lo contiene; es decir, es la intersección de todos los conjuntos convexos que lo contienen.

Demostración

Sea $\mathrm{Conv}(C)$ la envolvente convexa de un conjunto C; entonces:

i) Que $\mathrm{Conv}(C)$ es un conjunto convexo lo probaremos así: Sean $x = \sum_{i=0}^{p} r_i x_i$, $y = \sum_{i=0}^{q} t_i y_i$ donde $\sum_{i=0}^{p} r_i = \sum_{i=0}^{q} t_i = 1$; $r_i, t_i \geq 0$, $x_i, y_i \in C$, dos puntos de $\mathrm{Conv}(C)$, y tomamos $\lambda \in [0,1]$ cualquiera. Entonces,

$$\lambda x + (1-\lambda) y = \lambda \left(\sum_{i=0}^{p} r_i x_i \right) + (1-\lambda) \left(\sum_{i=0}^{q} t_i y_i \right)$$

$$= \sum_{i=0}^{p} (\lambda r_i) x_i + \sum_{i=0}^{q} ((1-\lambda) t_i) y_i \,,$$

y sólo restaría probar que $\sum_{i=0}^{p} \lambda r_i + \sum_{i=0}^{q} (1-\lambda) t_i = 1$; pero esto es cierto ya que

$$\sum_{i=0}^{p} \lambda r_i + \sum_{i=0}^{q} (1-\lambda) t_i = \lambda 1 + (1-\lambda) 1 = 1$$

ii) Ahora probemos que $\mathrm{Conv}(C)$ es la intersección de todos los conjuntos convexos que contienen a C:

a) Como cualquier conjunto convexo que contenga a C también contiene a todas las combinaciones lineales de puntos de C, entonces contiene a $\mathrm{Conv}(C)$. Por tanto, la intersección de todos los conjuntos convexos que contienen a C, contiene a $\mathrm{Conv}(C)$.

b) Ahora: $\text{Conv}(C)$ como conjunto convexo en sí mismo que contiene a C, tendrá que contener a la intersección de conjuntos convexos que contienen a C.

c) De a) y b) se deduce la igualdad de $\text{Conv}(C)$ con la intersección de todos los conjuntos convexos que contienen a C. ■

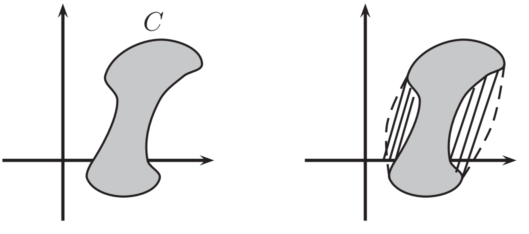


Figura 11. Conjunto C y su envolvente convexa

Ejercicios 1

1. ¿Por qué un conjunto de un sólo punto $C = \{x_0\}$ con $x_0 \in \mathbb{R}^n$, es un conjunto convexo?

2. Pruebe, utilizando la definición, que el conjunto $C = \{(x, y) \in \mathbb{R}^2 / y \geq x^3,\ x \geq 0\}$ es un conjunto convexo. Ilustre con una gráfica.

3. Pruebe que todo hiperplano en $\mathbb{R}^n$ es un conjunto convexo. Dé ejemplos específicos de hiperplanos en $\mathbb{R}^2$ y $\mathbb{R}^3$ y confirme su convexidad mediante un dibujo adecuado.

4. Pruebe que el conjunto $C = \{x \in \mathbb{R}^n / \parallel x \parallel = 1\}$ *no* es un conjunto convexo. ¿Cuál es su envolvente convexa?

5. Sólo por observación de su gráfica, confirme que los conjuntos siguientes son convexos:

   a) $\{(x, y) \in \mathbb{R}^2 / y \geq \dfrac{1}{x},\ x > 0\} \cap \{(x, y) \in \mathbb{R}^2 / y \geq x^2\}$

   b) $\{(x, y) \in \mathbb{R}^2 / x = 1\} \cap \{(x, y) \in \mathbb{R}^2 / xy + y^2 + 1 \geq 0\}$

   c) $3\{(x, y) \in \mathbb{R}^2 / x > 1, y > 3\} \cap 4\{(x, y) \in \mathbb{R}^2 / 3x + 2y \geq 7\}$

6. Determine la suma $C_1 + C_2$ en los siguientes casos y decida sobre su convexidad:

   a) $C_1 = \{(x, y) \in \mathbb{R}^2 / x^2 + y^2 \leq 1\}, \quad C_2 = \{(3, 4)\}$

b) $C_1 = \{(x, y) \in \mathbb{R}^2 / 2x - y = 3\}$, $C_2 = \{(x, y) \in \mathbb{R}^2 / x = y,\ 0 \leq x \leq 1\}$

c) $C_1 = [0, 1] \times [0, 1]$, $C_2 = [2, 3] \times [2, 3]$

d) $C_1 = \{(x, y) \in \mathbb{R}^2 / x - 5y = 1\}$, $C_2 = \{(-1, -1)\}$

[Indicación: un dibujo ayudaría].

7. Una prueba alterna de la afirmación de que todo hiperplano es convexo se basa en el ejercicio 1 de esta sección y en la parte b) del teorema 1, una vez se pruebe que todo subespacio vectorial de $\mathbb{R}^n$ es convexo. Demuestre esto último.

8. Pruebe que un conjunto convexo C es un cono convexo si, y sólo si, cualquier combinación lineal no-negativa de puntos de C es, de nuevo, un punto de C.

9. ¿Cuáles de los siguientes conjuntos son politopos, poliedros o conos convexos?:

a) $\{(x, y) \in \mathbb{R}^2 /\ x \geq 0,\ \ y \geq 0\}$

b) $\{(x, y) \in \mathbb{R}^2 /\ x \geq 0\}$

c) $\{(x, y) \in \mathbb{R}^2 /\ y + 5x \leq 0\}$

2. Introducción a la programación lineal

Ya habíamos resaltado el hecho de que una recta de la forma típica $ax+by+c = 0$ divide el plano cartesiano en dos *semiespacios* (o semiplanos) cerrados: uno, donde $ax+by+c \geq 0$ y, otro, donde $ax+by+c \leq 0$, con la recta como frontera entre los dos (figura 12).

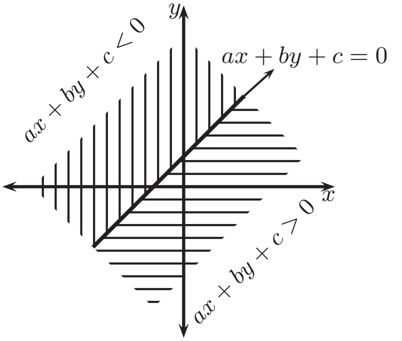


Figura 12

De manera similar, un plano con ecuación típica $ax + by + cz + d = 0$, divide el espacio tridimensional en dos semiespacios cerrados conformados por las triplas (x, y, z) tales que, o bien $ax + by + cz + d \geq 0$, o $ax + by + cz + d \leq 0$. La frontera de los semiespacios es el plano mismo.

Muchos problemas sobre optimización pueden plantearse en términos de semiespacios cerrados. En este tipo de problemas en particular, es corriente considerar variables no-negativas, lo cual implica estudiar los semiespacios cerrados $x_i \geq 0$, para $i = 1, \ldots, n$. Por ejemplo, considerar las tres condiciones $ax + by + c \geq 0$, $x \geq 0$ y $y \geq 0$, simultáneamente, nos conduce a una región como la dibujada en la figura 13 (dependiendo de si los coeficientes a, b, c son positivos, negativos o ceros) que corresponde al politopo de los tres semiespacios.

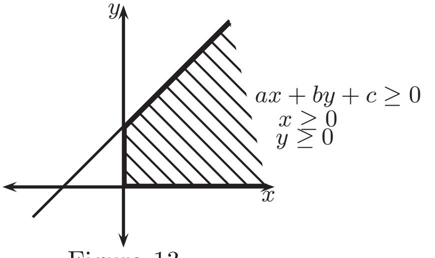


Figura 13

En numerosas ocasiones (y por razones del contexto en el que se estudia comúnmente) el nombre técnico usual de este politopo es *región factible* o *alcanzable* (*feasible set*). En este punto es interesante hacer una analogía recordando que cuando resolvíamos el sistema $Ax = b$ de m ecuaciones con n incógnitas, lo que realmente buscábamos era el politopo de la intersección de m hiperplanos en R^n.

Ahora: existe un campo de las matemáticas denominado *programación lineal* (ver nota 2 adelante), en donde se estudian problemas que buscan un valor óptimo de una cierta *función objetivo* lineal que está restringida a una cierta región factible definida a través de una serie de desigualdades. Así, un problema típico de *programación lineal* consiste en determinar la colección de puntos de la región factible, que maximizan o minimizan una cierta función objetivo lineal específica. Por ejemplo, podemos buscar el par (x, y), que satisfaciendo las desigualdades $x + 2y \geq 4$, $x \geq 0$ y $y \geq 0$, minimiza la función objetivo lineal $2x + 3y$. Este problema de optimización se acostumbra a escribir así:

$$\begin{array}{ll} \text{Minimizar} & 2x + 3y \\ \text{sujeta a} & x + 2y \geq 4 \\ & \quad\;\; x \geq 0 \\ & \quad\;\; y \geq 0 \end{array}$$

La figura 14 nos muestra la familia de rectas de nivel (Volumen 0 (Fundamentos)) que se construyen a partir de la función objetivo $2x + 3y = c$, para c variando sobre los números reales; así, nuestro problema se reduce a *encontrar la línea de más bajo nivel que interseca a la región factible.* Esta intersección ocurre, precisamente, en el punto A de la figura 14, donde $x^* = 0$, $y^* = 2$, y el mínimo valor de la función objetivo es $2(\,0\,) + 3(\,2\,) = 6$.

Como podemos ver, la solución al problema anterior ocurre en una esquina de la frontera del conjunto alcanzable. Pero esto no es una coincidencia: *es, de hecho, una característica esencial del comportamiento lineal de las funciones implicadas.*

Observemos, sin embargo, que en la misma figura 14, el problema

$$\begin{aligned} \text{Maximizar} \quad & 2x + 3y \\ \text{sujeta a} \quad & x + 2y \geq 4 \\ & x \geq 0 \\ & y \geq 0 \end{aligned}$$

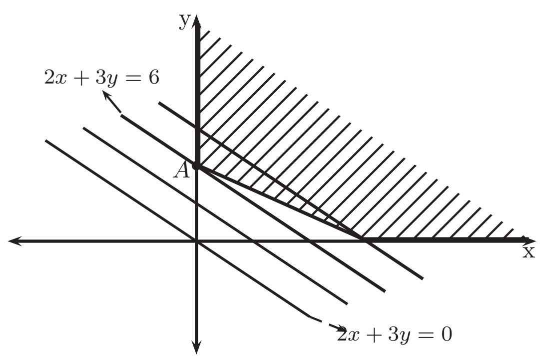


Figura 14. Un problema de programación lineal

no tiene solución: la función objetivo puede crecer arbitrariamente sobre el conjunto factible. No es muy difícil observar que en un problema de programación lineal general

$$\begin{aligned} \text{Maximizar} \quad & \sum_{i=1}^{n} c_i\, x_i \\ \text{sujeta a} \quad & Ax \geq b \\ & x \geq 0 \end{aligned}$$

(donde $c_i \in \mathbb{R}$, $x = (\,x_1, \ldots, x_n\,)^T \in \mathbb{R}^n$, $A \in \mathfrak{M}_{m\times n}$, $b \in \mathbb{R}^m$ y $x \geq 0$ significa que $x_i \geq 0$ para todo $i = 1, 2, \ldots, n$), si se da que el conjunto factible $\{x \in \mathbb{R}/Ax \geq 0\}$ no es vacío, entonces la función objetivo crece indefinidamente, o alcanza un único valor con uno o infinitos vectores x^* óptimos; pero si la solución x^* es única, entonces x^* se alcanza en una esquina de la frontera del conjunto factible. En general, en problemas significativos en la práctica, es esto último lo que realmente ocurre. En los Ejercicios 2 que presentamos adelante pedimos al lector que verifique esto último mediante una adecuada representación gráfica de cada uno de los ejercicios. Sobre los problemas implicados por la "optimización lineal" (mejor conocida como "programación lineal") regresaremos en la lección 2 del Volumen 3 (Optimización y dinámica).

Nota 2. (Sobre los orígenes de la programación lineal)
Aunque ya en 1939, los matemáticos rusos L. Kantorovich y W. Karush tenían algunos avances en problemas de optimización lineal, las ideas centrales de lo que hoy conocemos como *programación lineal* serían formuladas en 1947 por George Bernard Dantzig[1] en el estudio de ciertos problemas militares mientras trabajaba para la Fuerza Aérea de los Estados Unidos. La necesidad de organizar y despachar tropas y abastecimientos condujo a "programar" (de allí el término "programación" que lleva el método) horarios de entrenamiento, ofertas logísticas, y desplazamientos de tropas. Dantzig mecanizó esto introduciendo entonces estructuras lineales (y, por tanto, de convexidad) que condujeron a la técnica que hoy llamamos *programación lineal.*

Dantzig desarrollo allí el *"método simplex"* para resolver estos problemas (lección 2, Volumen 3 (Optimización y dinámica)) en el que utilizaba los primeros computadores conocidos hasta entonces, para llevar a cabo numerosos, rápidos y precisos cálculos. Durante casi un año, Dantzig y sus colegas estudiaron miles de situaciones tomadas de la experiencia de la Segunda Guerra Mundial, y mostraron que muchas de ellas podían (con cierta aproximación) convertirse al formato de la programación lineal, con excepción de aquellas en las que la no-convexidad era central y fundamental. En este proceso, observó que la representación geométrica de estos problemas conducía a conjuntos poliedrales y que las soluciones mejoraban cuando se movía de un vértice a otro de este conjunto convexo, alcanzando eventualmente la solución óptima. Sin embargo, Dantzig no utilizaba en aquel momento la noción de función objetivo y por eso creyó que el método simplex era ineficiente en extremo y que no valía la pena investigar más allá.

En el mismo año de 1947, Dantzig contactaría al economista Tjalling C. Koopmans (premio Nobel de economía en 1975) quien se interesaría en el modelo de programación lineal como herramienta conveniente en un problema central de la teoría económica: *la distribución eficiente de recursos.* Y también consultaría, a finales del año 1947, a John von Neumann acerca de los procedimientos de solución pues éste venía trabajando con técnicas similares sobre

[1] Aunque publicadas en 1949.

lo que daría en llamar *"teoría de juegos"* (ver el"contexto económico" al final de esta lección).

Pero, entretanto, Dantzig y su grupo del Pentágono (particularmente, Jack Laderman) continuaban probando su método simplex y encontraban que funcionaba realmente bien: al final, se mostró que *no* se requerían tantos pasos de ajuste de un vértice a otro, como Dantzig creyó inicialmente, y así, la rama de las matemáticas hoy conocida como *programación lineal* se recibiría como herramienta esencial a muchas aplicaciones prácticas en economía, agricultura, medicina, ciencias sociales, transporte, nutrición etc.

Ejercicios 2

1. Resolver los siguientes problemas de programación lineal *mediante representación gráfica:*

   a)
   $$\begin{aligned} \text{Maximizar } & 4x + y \\ \text{sujeta a } & x + 2y \leq 5 \\ & 3x + 2y \leq 4 \\ & x \geq 0 \\ & y \geq 0 \end{aligned}$$

   b)
   $$\begin{aligned} \text{Maximizar } & 25x + 2y \\ \text{sujeta a } & x + 2y \leq 8000 \\ & 3x + 2y \leq 9000 \\ & x \geq 0 \\ & y \geq 0 \end{aligned}$$

2. Resolver, si es posible,

   $$\begin{aligned} \text{Maximizar } & cx^T \\ \text{sujeta a } & Ax \geq b \\ & x \geq 0 \end{aligned}$$

   para $c = [-10, 2],\ A = \begin{bmatrix} 1 & 1 \\ -1 & -5 \end{bmatrix},\ x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},\ b = (3, 4)$

3. Pruebe que en el problema típico de programación lineal

   $$\begin{aligned} \text{Maximizar } & cx^T \\ \text{sujeta a } & Ax \geq b \\ & x \geq 0 \end{aligned}$$

   el conjunto factible $\{x \in \mathbb{R}^n / Ax \geq b\}$ es convexo.

4. En el problema

$$\begin{aligned} \text{Maximizar} \quad & x - y \\ \text{sujeta a} \quad & -x - y \leq r \\ & sx + y \leq 10 \\ & x \geq 0 \\ & y \geq 0 \end{aligned}$$

a) ¿Para qué valores de r y s existe solución única?
b) ¿Para cuáles valores de r y s existen soluciones múltiples?
c) ¿Existen valores de r y s para los que el conjunto factible es vacío?

3. Contexto económico

a. Sobre la noción de convexidad en economía

El concepto de *conjunto convexo* como herramienta utilizada en los métodos matemáticos para la economía, fue introducida por John von Neumann en su modelo de equilibrio general (1932) y de la teoría de juegos (1944). En ambos casos, la necesidad de este concepto surgió a partir de la presencia de desigualdades de funciones lineales específicas de varias *variables positivas*. En la teoría de la asignación de recursos, la noción de convexidad se ha utilizado, típicamente, para especificar hipótesis acerca de las posibilidades de producción de una firma, o de la forma de elegir de un consumidor. A partir de la misma definición de convexidad, es claro que ésta sólo puede utilizarse con cierto grado de aproximación cuando exista alguna interpretación posible de la *divisibilidad de recursos o mercancías bajo estudio* (figura 15). De otro lado, en la teoría de juegos se ha utilizado para describir el conjunto de eventos aleatorios posibles en la toma de decisiones interactivas.

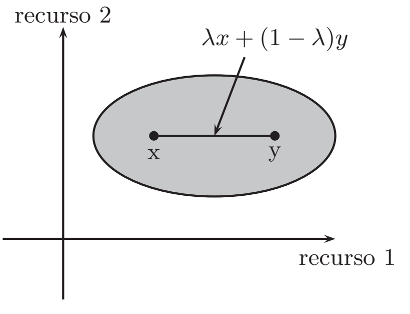


Figura 15. Si x, y son posibles, entonces $\lambda x + (1-\lambda)y$ también es posible, para cualquier $\lambda \in [0, 1]$

Una aplicación, particularmente relevante sobre la importancia del concepto de conjunto convexo en la teoría económica, apareció en el contexto de la *competencia perfecta* (es decir, cuando la influencia de cada agente sobre el resultado total de la actividad económica es insignificante). En 1964, el premio Nobel de Economía de 2005, Robert J. Aumann publicó *Markets with a Continuum of Traders* (Econometrica), en donde mostraba que si se asocia con cada agente de una economía en competencia perfecta, un conjunto arbitrario en el espacio de mercancías (es decir, un subconjunto de $\mathbb{R}^n_+$ donde n es el número de mercancías de la economía), y se "promedian" estos conjuntos sobre la colección de agentes, entonces el resultado es un *conjunto convexo*, dando así una justificación del *por qué se acostumbra asumir que los agentes de una economía competitiva siempre tienen sus conjuntos de consumo convexos.*

Pero, en general, está claro que el concepto de convexidad nos permite alcanzar hipótesis mínimas para la validez de una parte importante de los resultados matemáticos del núcleo de la teoría económica actual. Ayuda a simplificar y abreviar la formulación de los problemas y a utilizar las técnicas más simples disponibles con el propósito de esclarecer cuáles son algunas de las preguntas esenciales del modelo bajo discusión. Aún así, siempre deberíamos estar listos a responder el porqué de la hipótesis de convexidad en cada caso particular y en cada modelo específico, y es muy probable que muchas de las respuestas converjan alrededor de la idea de que *para este tipo de conjuntos, los resultados y teoremas matemáticos que se consideran convenientes, están particularmente bien adaptados a las necesidades del investigador, además de ser muy poderosos y agudos en sus conclusiones.*

b. Tres modelos lineales básicos de la teoría económica

En las tres subsecciones que siguen presentamos brevemente (y en su *versión lineal*) tres modelos básicos de la teoría económica actual. El primero, es el modelo macroeconómico "keynesiano" IS-LM de Hicks (1937); el segundo es el modelo de la teoría de juegos de von Neumann y Morgenstern (1944); y el tercero, el modelo lineal general de la teoría microeconómica tradicional: el análisis de actividades de Koopmans (1951).

I). El modelo "keynesiano" IS-LM lineal (Hicks (1937))

Como afirmábamos en el "contexto económico" de la lección 1, después de la publicación de la *Teoría general* de Keynes en 1936, muchos economistas intentaron expresar sus contenidos principales mediante ecuaciones, buscando aclarar las interrelaciones entre el "mercado monetario" y el "mercado real". Entre ellos estuvieron Roy Harrod (1937)[2], James Meade (1937)[3], Oskar Lange (1936)[4], John Hicks (1937)[5] y Alvin Hansen (1947, 1953)[6, 7]. Pero, de estos, ha sido el *Mr. Keynes and the Classics: A Suggested Interpretation* (1937) de John Hicks, el de mayor impacto y popularidad. Allí, Hicks dibujó dos curvas "SI-LL" para ilustrar estas interrelaciones, y desde entonces (ayudadas por los artículos de Hansen), estas curvas se han transformado en el instrumento de

[2] Harrod, Roy (1937), Mr. Keynes and Traditional Theory, *Econometrica*, vol. 42.
[3] Meade, J. (1937), A Simplified Model of Mr. Keynes System, *Review of Economic Studies*, vol. 4, 98-16.
[4] Lange, O. (1936), On the Economic Theory of Socialism, *Review of Economic Studies*, vol. 4 (1), 53-71.
[5] Hicks, J. (1937), Mr. Keynes and the Classics: A Suggested Simplification, *Econometrica*, vol. 5 (2), 147-159.
[6] Hansen, A. (1947), *Keynes on Economic Policy*, en Harris, editor, New Economics.
[7] Hansen, A. (1953), *A Guide to Keynes*, New York: Mc Graw Hill.

análisis económico conocido como el *modelo IS-LM*. Basados en las presentaciones originales de Hicks y Hansen, el mejor lugar para comenzar quizás sea el conocido diagrama ingreso-gasto o "cruz keynesiana" que Paul Samuelson (1948)[8], Abba Lerner (1951,1952)[9,10] y el mismo Hansen (1953) ayudaran a popularizar.

Cabe advertir que la siguiente breve construcción *lineal* del modelo IS-LM, basada en el modelo original de Hicks (1937), ignora algunos problemas fundamentales de la teoría "keynesiana" que más adelante destacaremos (Volumen 3 (Optimización y dinámica)).

La curva IS lineal

En una economía, la curva IS lineal se describe mediante las combinaciones en el mercado de bienes entre los tipos de interés (i), los niveles de ingreso (renta) agregado (Y), y el gasto autónomo (o índice de la política fiscal (F)), de la siguiente forma:

$$i = \frac{1}{\beta}\,[\,F - (\,1 - c\,)Y\,] \qquad \text{(IS)}$$

donde c es la propensión marginal a consumir ($0 < c < 1$) y $\beta > 0$ es la inversión marginal (con respecto al tipo de interés).

La curva LM lineal

Esta curva describe, mediante una función lineal, las combinaciones en el mercado monetario entre los tipos de interés (i); los saldos reales $\left(\frac{M}{P}\right)$ y los niveles de ingreso (renta) agregado(Y):

$$\frac{M}{P} = kY - hi$$

y así,

$$i = \frac{1}{h}\left[kY - \frac{M}{P}\right] \qquad \text{(LM)}$$

donde M son los saldos monetarios nominales; P es el nivel de precios; y donde $k,\ h > 0$ representan las demandas marginales de saldos reales con respecto a la renta y al tipo de interés, respectivamente.

[8] Samuelson, P. (1948), International Trade and the Equalisation of Factor Prices, *Economic Journal*, vol. 58 (230), 163-184.
[9] Lerner, A. (1951), *The Economics of Employment*, New York: Mc Graw Hill.
[10] Lerner, A. (1952), Factor Prices and International Trade, *Economica*, vol. 19, 67-84.

Equilibrio

Tenemos entonces que el sistema IS-LM es un sistema de dos ecuaciones lineales con dos incógnitas:

$$i = \frac{1}{\beta}\,[\,F - (\,1 - c\,)Y\,] \qquad \text{(IS)}$$
$$i = \frac{1}{h}\left[kY - \frac{M}{P}\right] \qquad \text{(LM)}$$

Inicialmente esta curva fue llamada por Hicks "la curva L", pero Hansen la llamó "curva LM" enfatizando el hecho de que representaba una curva a lo largo de la cual L, que es la demanda de dinero, igualaba la oferta de dinero M. En un diagrama Y *vs* i se obtiene la figura 16.

A partir de la situación de equilibrio que resulta de resolver, simultáneamente, las ecuaciones (IS) y (LM), es decir, de

$$Y^* = \frac{hF + \beta\,\frac{M}{P}}{\beta k + h(\,1 - c\,)} \qquad\qquad i^* = \frac{kF - (\,1 - c\,)\frac{M}{P}}{\beta k + h(\,1 - c\,)},$$

podemos estudiar diversas situaciones hipotéticas:

a) Si $h = 0$ (es decir, la renta es indiferente a la política fiscal) se obtiene que

$$Y^* = \frac{1}{k}\frac{M}{P}, \qquad i^* = \frac{kF - (\,1 - c\,)\frac{M}{P}}{\beta k} \qquad \text{(Caso “clásico”: } LM \text{ es vertical)}$$

b) Si h crece indefinidamente, entonces

$$Y^* = \frac{F}{1 - c}, \qquad i^* = 0 \qquad \text{(“trampa de liquidez”: LM es horizontal)}$$

Así, sólo la política fiscal determina el nivel de renta, y los tipos de interés son nulos. En la trampa de liquidez el dinero no importa para el cálculo de la renta.

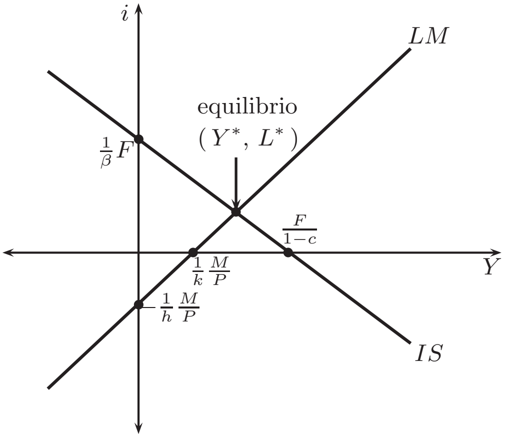


Figura 16. La cruz keynesiana

Nota 3.

Muchos keynesianos, tales como Luigi Pasinetti (1974)[11], afirman que el sistema de Keynes no debería pensarse (y este es el caso del modelo IS-LM) como la *solución simultánea* de la interacción entre los mercados reales y monetario, sino *secuencialmente* o mediante "bloques recursivos". Más específicamente, afirma que el sistema keynesiano debería verse como una sucesión de decisiones alternantes entre el "mercado financiero" y el "mercado de bienes", determinándose primero la tasa de interés mediante una decisión de portafolio (curva LM) en los mercados financieros y, sólo después, determinar la inversión, la producción y el empleo en el mercado real (curva IS) el cual, entonces, retroalimenta otra decisión de portafolio, etc. En el fondo, el problema aquí es que la solución simultánea del sistema elimina los conceptos que dependen del tiempo y que algunos otros economistas como Richard Kuhn (1984) y Joan Robinson (1973, 1978, 1979)[12, 13, 14] consideran fundamentales a la teoría keynesiana; en particular, ignora la incertidumbre, las expectativas, la especulación (es decir, los "espíritus animales" como Keynes los llamaba). Inclusive, posteriormente, el propio Hicks (1980, 1988)[15, 16] reconocía que la simultaneidad del modelo IS-LM lo hace incongruente.

Las muchas "historias económicas reales" que el diagrama IS-LM de Hicks y

[11] Pasinetti, L. (1974), *Growth and Income Distribution: Essays in Economic Theory*, Cambridge: Cambridge University Press.

[12] Robinson, J. (1973), *Collected Economic Papers*, Vol. IV, Oxford: Basil Blackwell.

[13] Robinson, J. (1978), Keynes and Ricardo, *Journal of Poskeynesian Economics*, vol. 36 (1), 31-58.

[14] Robinson, J. (1979), *The Generalization of the General Theory and Other Essays*, London: MacMillan.

[15] Hicks, J. (1980), IS-LM: An Explanation, *Journal of Poskeynesian Economics*, vol. 3, 139-155.

[16] Hicks, J. (1988), *Towards a More General Theory*, en *Finance Constraints, Expectations and Macroeconomics*, ed. Kohn y Tsiang.

Hansen generan, a veces no permite observar las dificultades teóricas y lógicas que le subyacen. Sin embargo, es claro a partir del trabajo de economistas keynesianos como Abba Lerner (1944, 1951, 1952)[17, 18, 19], Tibor Scitovsky (1940)[20], Sidney Weintraub (1958, 1959, 1961, 1966)[21, 22, 23, 24] y Paul Davidson (1972, 1994)[25, 26] (quienes nunca utilizaron el aparato IS-LM), que este no es ni el único, ni el más creíble, ni siquiera el más coherente para expresar la *Teoría general* de Keynes, pero que sí puede considerarse, tal vez, el más simple. El modelo IS-LM general lo estudiaremos en la lección 3 del volumen 3 (Optimización y dinámica).

II). La teoría de juegos de Von Neumann y Morgenstern (1944)

En 1928, John von Neumann reportó a la Sociedad Matemática de Göttingen que había conseguido encontrar una "estrategia racional" al problema de lanzar una moneda al aire[27]. Y aunque esto no pareciera ser un logro de mayores perspectivas, fue el comienzo de lo que hoy conocemos como *teoría de juegos*.

La prueba de von Neumann mostraba que existía un "mejor método posible" de juego que era matemáticamente determinable, y la "mejor estrategia posible" era aquella que le *aseguraba a un jugador la máxima ventaja sin importar lo que los oponentes hicieran*. La clave aquí fue que existiera tal "estrategia óptima" y que el juego entonces tuviera "solución". Von Neumann entonces se convenció de que la solución a ciertos problemas de juegos podría arrojar luz sobre algunos problemas económicos y de confrontación. Y así fue: en la época de la posguerra, la teoría de juegos fue más allá del reino de lo económico y pasó a la geopolítica.

[17] Lerner, A. (1944), Interest Theory: Supply and Demand for Loans or Supply and Demand for Cash?, *Review of Economic Studies*, vol. 26, 88-91.

[18] Lerner, A. (1951), *The Economics of Employment*, New York: Mc Graw Hill

[19] Lerner, A. (1952), The Essential Properties of Interest and Money, *Quarterly Journal of Economics*, vol. 66, 172-193.

[20] Scitovsky, T. (1940), A Study of Interest and Capital, *Economica*, vol. 7, 293-317.

[21] Weintraub, S. (1958), *An Approach to the Theory of Income Distribution*, Philadelphia: Chilton.

[22] Weintraub, S. (1959), *A General Theory of the Price Level, Output and Income Distribution*, Philadelphia: Chilton.

[23] Weintraub, S. (1961), *Classical Keynesianism: A Plea for its Abandonnment*, Philadelphia: Chilton.

[24] Weintraub, S. (1966), *A Keynesian Theory of Employment, Growth and Income Distribution*, Philadelphia: Chilton.

[25] Davidson, P. (1972), A Keynesian View of Friedman's Theoretical Framework for Monetary Analysis, *Journal of Political Economy*, vol. 80 (5), 864-882.

[26] Davidson, P. (1994), *The Asimakopulos View of Keynes's General Theory* in *Investment and Employment in Theory and Practice*, ed. Harcourt and Roncaglia.

[27] La presente sección está basada en el capítulo 1 del texto *Un Curso de Teoría de Juegos Clásica* (2005), (Sergio Monsalve y Julián Arévalo (Editores)), Universidad Externado de Colombia.

En las primeras épocas de la teoría de juegos después de la aparición en 1944 del monumental y fundacional trabajo de von Neumann y Morgenstern, *Theory of Games and Economic Behavior*, se creyó que el valor de la contribución de la nueva técnica a los problemas económicos era dudoso. Y la razón de esto es que la mayoría de las situaciones económicas que estudiaron von Neumann y Morgenstern toman la forma de juegos en los cuales, o bien la suma de las ganancias de los actores involucrados no es constante (juegos de suma no-constante), o el número de actores es mayor que 2, y en estos casos más complicados la teoría de von Neumann y Morgenstern no es tan contundente. Sin embargo, el futuro le depararía a la teoría de juegos, extensiones que han hecho de ella una herramienta fundamental para el análisis económico y social de hoy, como veremos más adelante (Volumen 3 (Optimización y dinámica), lección 3).

Un bosquejo de los aportes básicos de von Neumann y Morgenstern (1944)

Todos los juegos que von Neumann y Morgenstern estudian tienen varios elementos en común:

a) Un conjunto *finito* de jugadores (que pueden ser personas, animales o "entidades" aún más abstractas) y cada jugador tiene a su disposición un conjunto *finito* de reglas (o estrategias) para jugar.

b) Cada jugador tiene *conocimiento completo* de las reglas del juego y de los jugadores.

c) El juego termina después de un número finito de etapas.

d) Después de que el juego termina, se le asigna un pago numérico (que es positivo si ha ganado en el juego y negativo si se ha perdido) que a su vez es una suma ponderada de los pagos recibidos en cada etapa.

e) Existen posibles "movimientos" de *Natura*; es decir, se permiten ciertas formas de incertidumbre en las jugadas de los jugadores.

Y dieron en clasificar estos elementos con cuatro criterios: número de jugadores, estrategias, características de los pagos, y acuerdos *antes* de comenzar el juego. Esta clasificación marcaría el derrotero de la teoría de juegos durante varias décadas.

Juegos de dos jugadores y de suma cero

En este tipo de juego, que es quizás el más simple que podemos encontrar, tenemos dos jugadores, 1 y 2, con respectivos conjuntos finitos de estrategias a su disposición, C_1 y C_2; y también con funciones de pagos asociadas, π_1 y π_2, que dependen no sólo de su elección particular sino también de la elección del otro; es decir, π_1 y π_2 son funciones con dominio $C_1 \times C_2$ (el producto cartesiano de C_1 y C_2) y recorrido en los números reales. Toda esta información la resumimos en la siguiente bimatriz:

$$
\begin{array}{cccccc}
 & & & \text{Jugador 2} & & \\
 & & 1 & 2 & \cdots & \text{n} \\
 & 1 & \pi^1_{11}, \pi^2_{11} & \pi^1_{12}, \pi^2_{12} & \cdots & \pi^1_{1n}, \pi^2_{1n} \\
 & 2 & \pi^1_{21}, \pi^2_{21} & \pi^1_{22}, \pi^2_{22} & \cdots & \pi^1_{2n}, \pi^2_{2n} \\
 & 3 & \pi^1_{31}, \pi^2_{31} & \pi^1_{32}, \pi^2_{32} & \cdots & \pi^1_{3n}, \pi^2_{3n} \\
\text{Jugador 1} & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & \text{i} & \pi^1_{i1}, \pi^2_{i1} & \pi^1_{i2}, \pi^2_{i2} & \cdots & \pi^1_{in}, \pi^2_{in} \\
 & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & \text{m} & \pi^1_{m1}, \pi^2_{m1} & \pi^1_{m2}, \pi^2_{m2} & \cdots & \pi^1_{mn}, \pi^2_{mn}
\end{array}
$$

donde $C_1 \equiv \{1, 2, \ldots, m\}$, $C_2 \equiv \{1, 2, \ldots, n\}$; π^1_{ij} es el pago al jugador 1 cuando juega la estrategia i y su oponente juega j, y π^2_{ij} es el pago al jugador 2 cuando juega la estrategia j y su oponente juega i.

Pero si además asumimos que

$$\pi^1_{ij} + \pi^2_{ij} = 0 \qquad \text{(juego de suma cero)}$$

para todo $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$, entonces $\pi^1_{ij} = -\pi^2_{ij}$ y la descripción del juego dada por la tabla anterior ahora se simplifica:

$$
\begin{array}{cccccc}
 & & & \text{Jugador 2} & & \\
 & & 1 & 2 & \cdots & \text{n} \\
 & 1 & \pi_{11} & \pi_{12} & \cdots & \pi_{1n} \\
 & 2 & \pi_{21} & \pi_{22} & \cdots & \pi_{2n} \\
 & 3 & \pi_{31} & \pi_{32} & \cdots & \pi_{3n} \\
\text{Jugador 1} & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & \text{i} & \pi_{i1} & \pi_{i2} & \cdots & \pi_{in} \\
 & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & \text{m} & \pi_{m1} & \pi_{m2} & \cdots & \pi_{mn}
\end{array}
$$

donde $\pi_{ij} \equiv \pi^1_{ij} = -\pi^2_{ij}$.

El juego entonces consiste en que el jugador 1 escoge "entre filas", y el jugador 2 (simultáneamente) escoge "entre columnas", ambos buscando *maximizar* sus pagos (jugadores racionales). Von Neumann y Morgenstern consideraban que los jugadores (racionales como eran) elegirían de acuerdo con dos reglas particulares:

a) Dado que es un juego de suma cero, el jugador 1 (jugador fila) escogería la estrategia i que le maximizará los pagos y que a su vez le minimizará los de su oponente (maximizando ganancias), es decir, resolvería

$$\max_i \min_j \ \pi_{ij}$$

encontrando lo que llaman la *estrategia de maxmin* y la denotaban por v_1 = ventaja que el jugador 1 obtiene por jugar el juego.

b) De manera similar, el jugador 2 (jugador columna) *sabiendo* que su oponente seleccionará la fila con máximo pago[28], tratará de minimizar esto con una adecuada elección de su columna (es decir, minimiza pérdidas) resolviendo

$$\min_j \max_i \ \pi_{ij}$$

hallando lo que denominan la *estrategia de minmax* y la denotaban v_2 = ventaja que el jugador 2 obtiene por jugar el juego.

Es claro que v_1 y v_2 pueden ser diferentes. Sin embargo, von Neumann y Morgenstern encuentran que una posible *solución consistente* al juego es aquella estrategia $(\, i, j\,)$ que satisface las dos condiciones: maximizar las ganancias igual a minimizar las pérdidas; es decir,

$$v_1 = \max_i \min_j \ \pi_{ij} = \min_j \max_i \ \pi_{ij} = v_2$$

A este valor lo llamaron un *punto (de equilibrio) de silla*[29] del juego o, simplemente, el *valor del juego*.

Veamos unos cuantos ejemplos que ilustren esto.

---

[28] Esta es la hipótesis de *conocimiento común* de la racionalidad por parte de los jugadores.

[29] ¿Podría el lector justificar este nombre?

Ejemplo 6. (Juego de *tirar la moneda*)
El juego de *tirar la moneda* (*matching pennies*) consiste en dos jugadores que, simultáneamente, lanzan una moneda. Si en ambas monedas aparece *cara*, o en ambas aparece *sello*, el jugador 1 gana; pero si en una moneda aparece *cara* y en la otra *sello*, será el jugador 2 el que gana. Este juego es de dos jugadores y suma cero que puede escribirse mediante la siguiente matriz:

		Jugador 2	
		cara	sello
Jugador 1	cara	1	–1
	sello	–1	1

Aquí, $\pi_{11} = 1$, $\pi_{12} = -1$, $\pi_{21} = -1$, $\pi_{22} = 1$, y

$$v_1 = \max_i \min_j \ \pi_{ij} = \max\{\, \pi_{12} = -1, \pi_{21} = -1 \,\} = -1$$
$$v_2 = \min_j \max_i \ \pi_{ij} = \min\{\, \pi_{11} = 1, \pi_{22} = 1 \,\} = 1$$

Obviamente, $v_1 \neq v_2$ y no existe un punto de silla (o solución consistente) para este juego. Más adelante veremos cómo von Neumann y Morgenstern resuelven este problema. ▲

Ejemplo 7. (Juego de *piedra-papel-tijera*)
Este es el conocido juego infantil *piedra-papel-tijera* propuesto por von Neumann y Morgenstern, en el que "piedra vence a tijera", "tijera vence a papel" y "papel vence a piedra", y es un empate en los otros casos. Podemos describir este juego en una matriz como la siguiente:

			Jugador 2	
		piedra	papel	tijera
	piedra	0	−1	1
Jugador 1	papel	1	0	−1
	tijera	−1	1	0

Aquí,

$$v_1 = \max_i \min_j \ \pi_{ij} = \max\{\, \pi_{12} = -1, \pi_{23} = -1, \pi_{31} = -1 \,\} = -1$$
$$v_2 = \min_j \max_i \ \pi_{ij} = \min\{\, \pi_{21} = 1, \pi_{32} = 1, \pi_{13} = 1 \,\} = 1$$

También en este caso v_1 y v_2 son diferentes, y no existe un punto de silla. Nuevamente, adelante veremos cómo se "resuelve" este problema. ▲

Lo sucedido en los ejemplos clásicos de *tirar la moneda* y *piedra-papel-tijera* obligó a von Neumann y Morgenstern a confrontar el problema de encontrar una solución general a todos los juegos de dos personas y suma cero, y para ello utilizaron la teoría de las probabilidades. *Aparece, así, por primera vez en la teoría económica, el concepto de estrategia mixta* que no es más que una combinación probabilística de las estrategias *puras* hasta ahora estudiadas.

Una *estrategia mixta para el jugador* 1 es un vector de probabilidades $p = (p_1, p_2, \ldots, p_m)$, donde p_i $(i = 1, 2, \ldots, m)$ es la probabilidad de que el jugador 1 juegue la estrategia i; obviamente, asumimos que $p_i \geq 0$ y $\sum_{i=1}^{m} p_i = 1$. De manera similar, una *estrategia mixta para el jugador* 2 es un vector de probabilidades $q = (q_1, q_2, \ldots, q_n)$, donde q_j $(j = 1, 2, \ldots, n)$ es la probabilidad de que el jugador 2 juegue la estrategia j; asumiremos también que $q_j \geq 0$ y $\sum_{j=1}^{n} q_j = 1$. Notemos que las estrategias puras pueden verse como casos particulares de las mixtas; así, por ejemplo, $(0, 0, 1, 0, \ldots, 0)$ es la representación de la tercera estrategia por parte de alguno de los jugadores.

En este punto los autores utilizan el concepto de *valor esperado* probabilístico de una estrategia mixta, que es, simplemente, una valoración (ponderada por las probabilidades) de los pagos que recibiría en cada una de las estrategias puras. El pago esperado por el jugador 1 bajo la distribución de probabilidades $p = (p_1, \ldots, p_m)$ si el jugador 2 juega la estrategia j es

$$E(p, j) = p_1\pi_{1j} + p_2\pi_{2j} + \cdots + p_m\pi_{mj}$$

Se espera también aquí que el jugador 1 escoja las probabilidades p de tal forma que resuelva

$$\max_p \min_j E(p, j)$$

De manera similar, el jugador 2 recibiría un pago esperado bajo la distribución $q = (q_1, \ldots, q_n)$ si el jugador 1 juega la estrategia i igual a

$$E(i, q) = q_1\pi_{i1} + q_2\pi_{i2} + \cdots + q_n\pi_{in};$$

y se espera entonces que el jugador 2 escoja las probabilidades q de tal forma que resuelva

$$\min_q \max_i E(i, q)$$

Si p y q son tales que

$$\max_p \min_j E(p, j) = \min_q \max_i E(i, q)$$

Von Neumann y Morgenstern afirman entonces que estas probabilidades son una *solución al juego* (o un *punto de silla* del juego) y a este valor lo llaman el *valor del juego*.

Con estas herramientas de solución, regresemos a los juegos anteriores para ver si podemos ahora "resolverlos".

Ejemplo 8. (Juego de *tirar la moneda*, otra vez)
La justificación de las estrategias mixtas para este juego la dan von Neumann y Morgenstern de la siguiente manera: puesto que ninguna forma particular de "jugar" (*cara* o *sello*) es mejor que otra, y si todo lo que importa es averiguar las intenciones del oponente, no tendremos manera de encontrar una solución. Pero si el jugador no sólo intenta averiguar lo que el otro jugador va a mover sino que también se concentra en que no descubran sus intenciones, jugar "irregularmente" *cara* y *sello* podría ser una estrategia conveniente. Esto último es lo que se presenta como "$\frac{1}{2}$ de probabilidad de jugar *cara*; y $\frac{1}{2}$ de probabilidad de jugar *sello*". El punto aquí es que este procedimiento "protege" de pérdidas. De todas formas, el pago esperado de jugar esta estrategia es 0 para ambos.

Para ver esto, dibujemos el primer cuadrante donde el eje de abscisas está determinado por la probabilidad p, y el eje de ordenadas está determinado por el pago esperado $E(p)$ del jugador 1.

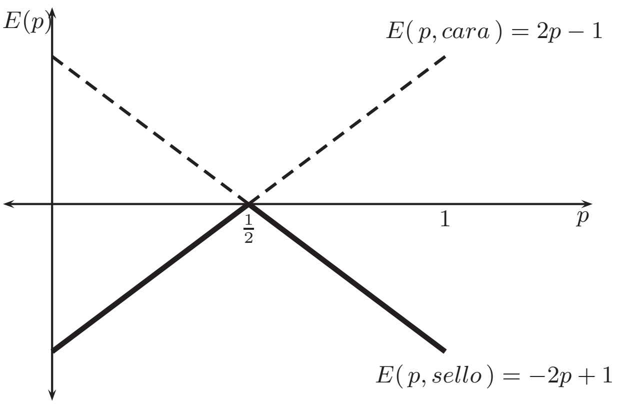


Figura 17. Solución a "tirar la moneda"

Observemos que para el jugador 1,

$$E(p, cara) = p(1) + (1 - p)(-1) = 2p - 1$$
$$E(p, sello) = p(-1) + (1 - p)(1) = -2p + 1$$

La línea resaltada de la figura 17, formada por los dos segmentos, es la gráfica de la función

$$E(p) = \text{mín}\{\, E(p, cara), E(p, sello)\,\} = \begin{cases} 2p-1 & \text{si} \quad p \leq \frac{1}{2} \\ -2p+1 & \text{si} \quad p \geq \frac{1}{2} \end{cases}$$

y el problema del jugador 1 es encontrar p que haga $E(p)$ lo máximo posible. Este valor ocurre en $p = \frac{1}{2}$ y $v_1 = 0$ que es el más alto pago esperado por el jugador 1, independientemente de lo que haga el jugador 2. El jugador 2 tiene un problema similar que se resuelve mediante el mismo tipo de análisis. Su solución muestra que, al igual que el jugador 1, puede dejarle su decisión a una moneda; es decir, adoptar $q = \frac{1}{2}$. Luego el valor del juego es $v_1 = v_2 = 0$, que se alcanza cuando $p = q = \frac{1}{2}$.

Ejemplo 9. (Juego de *piedra-papel-tijera*, otra vez)
La situación en *piedra-papel-tijera* es similar a la de *tirar la moneda*. El sentido común dice que la forma de jugar este juego es jugar las tres alternativas cada una con probabilidad $\frac{1}{3}$; y la teoría lo corrobora. Decíamos antes que la matriz de pagos en este caso era la siguiente:

			Jugador 2		
			$[q_1]$	$[q_2]$	
			piedra (Pi)	papel (Pa)	tijera (Ti)
	$[p_1]$	piedra (Pi)	0	-1	1
Jugador 1	$[p_2]$	papel (Pa)	1	0	-1
		tijera (Ti)	-1	1	0

En este caso, para el jugador 1, si $p = (p_1, p_2, 1-p_1-p_2)$ y $q = (q_1, q_2, 1-q_1-q_2)$ son las probabilidades de juego de los jugadores 1 y 2 , respectivamente, entonces:

$$\begin{aligned} E(p, Pi) &= p_1(0) + p_2(1) + (1-p_1-p_2)(-1) \\ E(p, Pa) &= p_1(-1) + p_2(0) + (1-p_1-p_2)(1) \\ E(p, Ti) &= p_1(1) + p_2(-1) + (1-p_1-p_2)(0) \end{aligned}$$

Por tanto, tenemos que:

$$\begin{aligned} E(p) &= \text{mín}\{\, E(p, Pi), E(p, Pa), E(p, Ti)\,\} \\ &= \text{mín}\{\, p_1 + 2p_2 - 1,\; -2p_1 - p_2 + 1,\; p_1 - p_2\,\} \end{aligned}$$

$$= \begin{cases} p_1 + 2p_2 - 1 & \text{si} \quad 0 \leq p_1;\, p_2 \leq \frac{1}{3};\ \text{ó si } \frac{1}{3} \leq p_1 \leq \frac{2}{3},\ 0 \leq p_2 \leq \frac{2}{3} - p_1 \\ -2p_1 - p_2 + 1 & \text{si} \quad \frac{1}{3} \leq p_1 \leq 1,\ \frac{2}{3} - p_1 \leq p_2 \leq 1 \\ p_1 - p_2 & \text{si} \quad 0 \leq p_1 \leq \frac{1}{3},\ \frac{2}{3} \leq p_2 \leq 1 \end{cases}$$

Pero en estas regiones, las tres funciones $p_1 + 2p_2 - 1$, $-2p_1 - p_2 + 1$, y $p_1 - p_2$ son negativas o cero. Luego para encontrar $\max E(p)$ debemos buscar dónde $E(p)$ se anula; y esto se logra igualando las tres funciones anteriores:

$$p_1 + 2p_2 - 1 = -2p_1 - p_2 + 1 = p_1 - p_2$$

De este sistema de ecuaciones con dos incógnitas se encuentra, fácilmente, que $p_1 = p_2 = \frac{1}{3}$; de manera similar para el jugador 2. Así, el valor de este juego es $v_1 = v_2 = 0$.

El teorema del minimax (von Neumann (1928))

Dados dos jugadores, 1 y 2, el primero con m estrategias y el segundo con n estrategias, y además el juego es de *suma cero* (lo que pierde un jugador lo gana el otro), a éste se acostumbra llamar un *juego de matriz (matrix game)* pues, obviamente, la descripción del juego es una matriz $m \times n$ de la forma

$$A = \begin{bmatrix} \pi_{11} & \cdots & \pi_{1n} \\ \vdots & \cdots & \vdots \\ \pi_{m1} & \cdots & \pi_{mn} \end{bmatrix}$$

donde la entrada π_{ij} representa el pago recibido por el jugador 1 cuando éste escoge la estrategia i y su oponente, el jugador 2, escoge la estrategia j. Aún así, la existencia de un punto de equilibrio (de silla) no es, en absoluto, obvia.

Si el jugador 1 coloca las probabilidades $p = (p_1, \ldots, p_m)$ sobre sus respectivas estrategias, y el jugador 2 coloca sobre sus respectivas estrategias las probabilidades $q = (q_1, \ldots, q_n)$, entonces el pago esperado por el jugador 1 si juega la estrategia i y su oponente juega la estrategia j es $\pi_{ij} p_i q_j$. Así que sin condicionamiento sobre las jugadas de los oponentes, el *pago esperado total* es $\sum_{j=1}^{n} \sum_{i=1}^{m} \pi_{ij}\, p_i\, q_j$. Pero esto puede escribirse más fácilmente en notación matricial: *es pAq^T lo que el jugador 1 busca maximizar y el jugador 2 minimizar.*

Por ejemplo, en el juego de *tirar la moneda*, se tiene que

$$A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

y así,

$$pAq^T = (\,-2 + 4q\,)p - 2q + 1$$

Si el jugador 1 quiere maximizar este valor controlando p, entonces haría lo siguiente:

a) Escoger $p = 1$ si $-2 + 4q > 0$; es decir, si $q > \frac{1}{2}$.

b) Escoger $p = 0$ si $-2 + 4q < 0$; es decir, si $q < \frac{1}{2}$.

c) Escoger cualquier p si $q = \frac{1}{2}$.

De manera similar, si el jugador 2 quiere minimizar el mismo valor $pAq^T = (\,-2 + 4p\,)q - 2p + 1$ haría lo siguiente:

a) Escoger $q = 1$ si $-2 + 4p < 0$; es decir, si $p < \frac{1}{2}$.

b) Escoger $q = 0$ si $-2 + 4p > 0$; es decir, si $p > \frac{1}{2}$.

c) Escoger cualquier q si $p = \frac{1}{2}$.

Si queremos alcanzar *ambos objetivos*, la solución es que los dos jugadores escojan $\frac{1}{2}$ como su probabilidad; es decir, $p^* = q^* = \frac{1}{2}$. El valor del juego es, efectivamente, $p^*Aq^{*T} = 0$.

Regresando al problema general, hemos entonces entendido que el jugador 1 ha garantizado que ganará al menos la cantidad

$$\max_p \min_q pAq^T$$

y no puede esperar ganar más; y el jugador 2 hace lo opuesto: escogerá de tal manera que no pierda más de

$$\min_q \max_p pAq^T$$

y no espera mejorar más esta situación. Luego si queremos asegurar que la cantidad que el jugador 1 busca ganar coincida con la que el jugador 2 está dispuesto a perder, debería probarse la existencia de p^* y q^* tales que resuelvan

$$\max_p \min_q pAq^T = \min_q \max_p pAq^T$$

La existencia de este *punto de silla* fue probada por Von Neumann en 1928 (16 años antes de publicar su *Theory of Games and Economic Behavior*) en un artículo que, en su momento, pasó desapercibido: *Zur Theorie der Gesellschaftspiele*, inicialmente publicado en el *Mathematische Annalen* y traducido al inglés en 1959 en *Contributions to the Theory of Games* (A. W. Tucker y D. Luce (ed.)). Este teorema, que será demostrado en el Volumen 3 (Optimización y dinámica), puede enunciarse de la siguiente forma:

Teorema 3. ***(Teorema minimax (von Neumann (1928)))***
Para cualquier matriz $A_{m\times n}$, existen vectores $p \in \mathbb{R}^m$ y $q \in \mathbb{R}^n$ tales que el minimax sobre todas las estrategias mixtas iguala al maximin; es decir,

$$\max_p \min_q pAq^T = \min_q \max_p pAq^T$$

Esta cantidad es el valor del juego. Si el máximo en el lado izquierdo se alcanza en p^* y el mínimo en el lado derecho se alcanza en q^*, entonces ninguno querrá cambiar su estrategia unilateralmente, es decir,

$$p^*Aq^T \leq p^*Aq^{*T} \leq pAq^{*T}$$

para todos los vectores de probabilidad p, q.

Algo notable es que en la prueba de von Neumann se utiliza la misma matemática que utilizaron los fundadores de la teoría de la programación lineal (en particular, Dantzig) desde 1947 en adelante, basados en las notas de von Neumann. Que realmente el problema del minimax es un problema de programación lineal se ve en el hecho de que puede plantearse, desde el punto de vista del jugador 1, como:

$$\begin{aligned}
\text{Maximizar } & v \\
\text{sujeta a } & v \leq p_1a_{11} + p_2a_{21} + \cdots + p_ma_{m1} \\
& v \leq p_1a_{12} + p_2a_{22} + \cdots + p_ma_{m2} \\
& \vdots \\
& v \leq p_1a_{1n} + p_2a_{2n} + \cdots + p_ma_{mn} \\
& p_1 + p_2 + \cdots + p_m = 1 \\
& p_1, p_2, \cdots, p_m \geq 0
\end{aligned} \tag{$*$}$$

Y si hacemos $a_{ij} \geq 0$ sumando un término C apropiado a cada una de las entradas de la matriz A, tendremos que es posible asumir que $v > 0$. Ahora: Sea $y_i = p_i/v$ para $i = 1, 2, ..., m$. Entonces $y_1+y_2+\cdots+y_m = \frac{p_1}{v}+\frac{p_2}{v}+\cdots+\frac{p_m}{v} = \frac{1}{v}(p_1 + p_2 + \cdots + p_m) = \frac{1}{v}$. Así, v es máximo si, y sólo si, $y_1 + y_2 + \cdots + y_m$ es mínimo. Por tanto, podemos plantear el problema $(*)$ como

$$\begin{aligned}
\text{Minimizar } & y_1 + y_2 + \cdots + y_m \\
\text{sujeta a } & y_1a_{11} + y_2a_{21} + \cdots + y_ma_{m1} \geq 1 \\
& \vdots \qquad \vdots \qquad \vdots \\
& y_1a_{1n} + y_2a_{2n} + \cdots + y_ma_{mn} \geq 1 \\
& y_1, y_2, \ldots, y_m \geq 0
\end{aligned} \tag{$**$}$$

Así, en el juego de *tirar la moneda*, donde

$$A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

se tendría, en primer lugar, el problema de programación lineal $(*)$

$$\begin{aligned} &\text{Maximizar } v \\ &\text{sujeta a} \quad p_1 - (1 - p_1) \geq v \\ &\qquad\qquad\; - p_1 + (1 - p_1) \geq v \end{aligned}$$

o, equivalentemente,

$$\begin{aligned} &\text{Maximizar} \quad v \\ &\text{sujeta a} \qquad 2p_1 - 1 \geq v \\ &\qquad\qquad\quad - 2p_1 + 1 \geq v \\ &\qquad\qquad\quad p_1 \geq 0 \end{aligned}$$

cuya solución, $p_1 = \frac{1}{2}$, $v = 0$, se ve en la figura 18 a).

Pero también, después de sumar $+2$ a cada una de las entradas de la matriz $\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ y obtener $\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, el problema se podría plantear, como en $(**)$, así:

$$\begin{aligned} &\text{Minimizar } y_1 + y_2 \\ &\text{sujeta a} \quad 3y_1 + y_2 \geq 1 \\ &\qquad\qquad\; y_1 + 3y_2 \geq 1 \\ &\qquad\qquad\; y_1, y_2 \geq 0 \end{aligned}$$

cuya solución es $y_1 = y_2 = \frac{1}{4}$, $v = \frac{1}{y_1+y_2} = 2$; y así $p_1 = p_2 = \frac{1}{4}v = \frac{1}{2}$. Como habíamos sumado $+2$ a la matriz original, el valor del juego original es $v-2 = 2-2 = 0$. El problema del jugador 2 es similar y se obtiene $q_1 = q_2 = \frac{1}{2}$ después de ver que $x_1 = x_2 = \frac{1}{4}$.

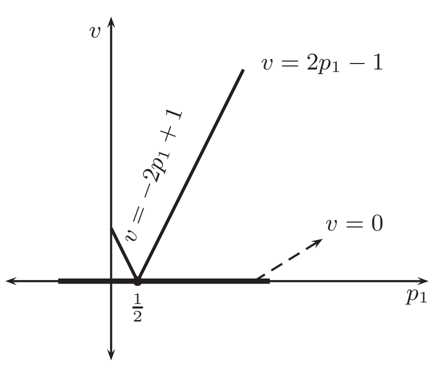


Figura 18 a)

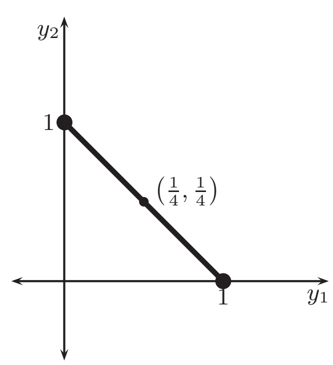


Figura 18 b)

Nota 4.
Sobre los modelos generales de la teoría de juegos clásica originada en los trabajos de von Neumann y Morgenstern, además de un estudio de la teoría general de interacciones, discutiremos en las lecciones 2 y 3 del Volumen 3 (Optimización y dinámica).

III). El modelo de análisis (lineal) de actividades (Koopmans (1951))

La idea del papel o de las ventajas de la utilización de precios para una mejor asignación de los recursos, ya sea a través del funcionamiento de mercados competitivos o como instrumentos para la planificación nacional, tiene una larga tradición en economía. En relación con los mercados competitivos, data por lo menos desde los tiempos de Adam Smith (1776)[30], y fue reformulada y desarrollada por el premio Nobel en economía de 1974, Friedrich von Hayek (1945)[31]. Entre los autores importantes que escribieron sobre la utilización de precios en la planificación socialista se cuentan Oskar Lange (1936)[32] y Abba Lerner (1937, 1938)[33, 34]. La novedad de los modelos de análisis de actividades consistió en el empleo de un *modelo lineal* y en el intento por desarrollar una *"teoría pre-institucional"* de la asignación de recursos. En las obras de Lange y Lerner se había sugerido que, en teoría, tanto la competencia perfecta como la planificación socialista implicaban una asignación eficiente de recursos (aunque en la realidad esto no fuera cierto). Y la idea era que después de mirar el problema pre-institucional, podría entonces procederse al diseño de instituciones que se aproximaran al modelo.

El *modelo lineal de análisis de actividades* del premio Nobel de economía de 1975, Tjalling C. Koopmans [1910–1985], es descendiente de la tradición de los modelos lineales de Walras-Cassel (equilibrio general), Wald (equilibrio general), von Neumann (crecimiento), Leontief (insumo-producto) de los cuales (excepto en los aspectos "dinámicos" del modelo de von Neumann) es generalización y síntesis. En él se describen las *tecnologías* mediante un conjunto de postulados (quizás los más sencillos posibles hasta tal momento), que permite estudiar el papel de los precios en la utilización eficiente de los recursos y, por tanto, desarrollar un papel fundamental dentro del modelo de equilibrio

[30] Smith, A. (1776), *An Inquiry into the Nature and Causes of the Wealth of Nations*, London: W. Strahan.

[31] Von Hayek, F. (1945), The Use of Knowledge in Society, *American Economic Review*, vol. 35 (4), 519-530.

[32] Lange, O. (1936),On the Economic Theory of Socialism, *Review of Economic Studies*, vol. 4 (1), 53-71.

[33] Lerner, A. (1937), Statics and Dynamics in Socialist Economics, *Economic Journal*, vo. 4, 72-76.

[34] Lerner, A. (1938), Theory and Practice of Socialist Economics, *Review of Economic Studies*, vol. 6 (1), 71-75.

general walrasiano. Veamos entonces en qué consiste el modelo de análisis de actividades.

Hipótesis del modelo de análisis de actividades

Puesto que el modelo (lineal) de análisis de actividades de Koopmans (1951)[35] es una representación de la actividad económica desde una visión walrasiana, entonces se presenta la típica división metodológica entre consumidores y productores, donde cada sector está descrito mediante ciertos *postulados* (o axiomas) que, en palabras del propio Koopmans, "delimitan un universo de discurso lógico en el cual el único criterio de validez es el de la implicación partiendo de ellos".

Postulado 1 (Sobre los agentes de decisión)

Existe un número dado de agentes que pueden subdividirse en l consumidores, m productores y p poseedores de recursos. Existe, también, un número finito n de bienes, subdivididos entre tipos de trabajos y otros bienes. Cada agente toma una decisión (para el período predeterminado) que consiste en la elección de una cantidad de cada tipo de trabajo y de cada bien; es decir, de un punto en $\mathbb{R}^n$. Se asume que los conjuntos de planes de consumo son *convexos* (figura 20), lo que necesariamente implica *divisibilidad perfecta* de los bienes y de los tipos de trabajo.

Postulado 2 (Sobre los consumidores)

El punto elegido por el i-ésimo consumidor debe pertenecer a un conjunto de consumo para $i = 1, \ldots, l$, cuyos puntos tienen coordenadas no-negativas para todos los bienes distintos al trabajo. Para las cantidades de los distintos trabajos ofrecidos escogemos coordenadas negativas. Si una cantidad de estas es cero, ello indicará que el bien no es consumido o que el tipo de trabajo no es ofrecido.

En la figura 20 el plan de consumo ofrece x_1 unidades de trabajo y consume x_2 unidades del bien 1.

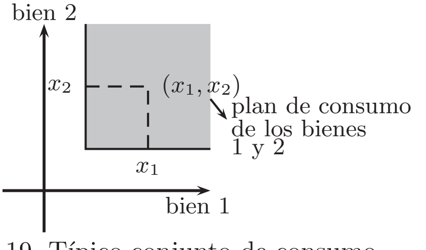


Figura 19. Típico conjunto de consumo

[35] Koopmans, T. C. (1951), *Activity Analysis of Production and Allocation*, New York: Wiley.

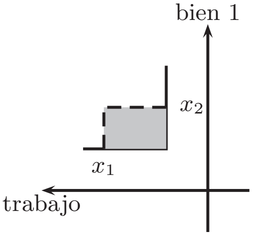


Figura 20. Plan de consumo

Sobre este conjunto existe, además, un *orden de preferencia completo* para el i-ésimo consumidor; es decir, sobre este conjunto se define un preorden completo "$\preceq$" ("menos preferido o indiferente a")[36] sobre los planes de consumo que describe la forma en que el consumidor elige entre las distintas alternativas de consumo. Se asume, además, que dado el plan de consumo $(x_1^*, x_2^*, .., x_n^*)$, el conjunto de planes de consumo $(x_1, x_2, .., x_n)$ tales que $(x_1, x_2, .., x_n) \succcurlyeq (x_1^*, x_2^*, .., x_n^*)$ es convexo; es decir, que un consumidor prefiere la *combinación* de las canastas, a una cualquiera de ellas (figura 22). En esta figura, es mejor la combinación $\lambda x_1 + (1-\lambda)x_2$ (con $\lambda \in (0,1)$) que x_1 ó x_2.

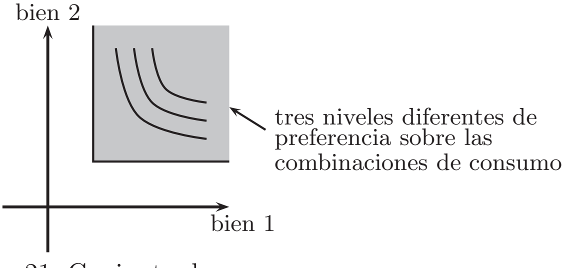


Figura 21. Conjunto de consumo
y orden de preferencias

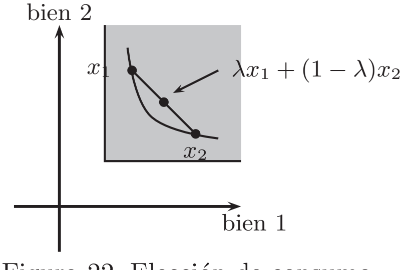


Figura 22. Elección de consumo

Postulado 3 (Sobre los productores)

El punto elegido por el j-ésimo productor debe pertenecer a un conjunto de producción para $j = 1, \ldots, m$, cuyos puntos tienen una coordenada *no-positiva*

[36] Recordemos que un preorden completo es una relación reflexiva, antisimétrica y transitiva (volumen 0 (Fundamentos)).

para cada tipo de trabajo y para cada insumo, además de coordenada *positiva* para cada tipo de bien producido. Este conjunto es *independiente* de las elecciones de los demás agentes. En la figura 23a aparece un politopo como conjunto de producción con dos insumos: insumo 1 y trabajo. En la figura 23b aparecen las correspondientes curvas de nivel en el plano trabajo vs insumo 1.

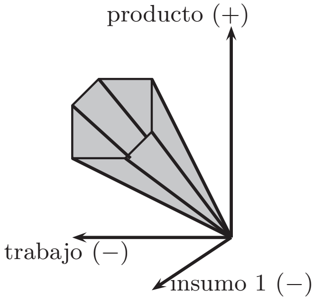


Figura 23a. Típico conjunto de producción

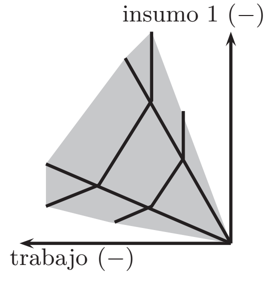


Figura 23b. Curvas de nivel de producción

Postulado 4 (Sobre los poseedores de recursos)

Cada poseedor de recursos controla una cantidad no-negativa de cada bien que no sea un tipo de trabajo, y elige desprenderse de una cantidad no-negativa de dichos bienes que no exceda las cantidades que posee.

Postulado 5 (No saturación local)

Dada una distancia positiva cualquiera, por pequeña que sea, para cada punto del conjunto de consumo de cualquier consumidor, existirá otro punto del mismo conjunto, preferido al primero, y cuya distancia a él es menor que la distancia dada (figura 24).

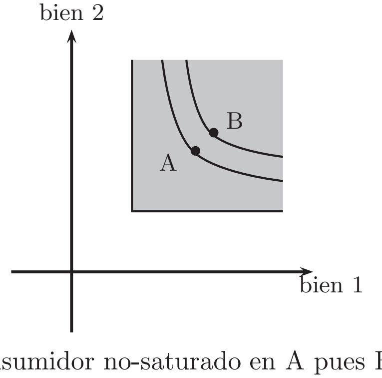


Figura 24. Consumidor no-saturado en A pues B es mejor

Este postulado implica que ningún consumidor podrá llegar a un estado máximo de saturación (saciedad) de todos los bienes disponibles en la economía.

Postulado 6 (Sobre los bienes)

Existe un número finito n de bienes, clasificados (sin yuxtaposición) entre l bienes deseados, p primarios, y $n - l - p$ intermedios. Cada bien puede existir en cualquier cantidad no-negativa en la que pueda producirse u obtenerse de la naturaleza. La conjunción o separación de cantidades de un mismo bien puede representarse mediante la suma y la diferencia de los números que miden dichas cantidades. Aquí, los *bienes deseados* representan bienes y servicios cuyo consumo o disponibilidad constituye el fin reconocido de la producción; los *bienes primarios* son los que se extraen directamente de la naturaleza; y los *bienes intermedios* son aquellos que simplemente pasan de un estado de producción a otro sin ser ni deseados por sí mismos, ni encontrarse disponibles en la naturaleza. El término "naturaleza" designa la fuente de bienes que permanece por fuera del sistema productivo que se estudia.

Postulado 7 (Existencia de actividades básicas de producción)

Existe un número finito m de actividades *básicas* de producción. Una actividad básica viene caracterizada por una cifra de producción neta para cada bien. En la figura 25 se muestran dos tipos de actividades básicas cuyo único insumo es el trabajo.

Estas actividades básicas son los métodos básicos de producción de la economía, y representan todo el conocimiento técnico básico disponible al comenzar el período de estudio.

Postulado 8 (Sobre la aditividad)

Dadas dos actividades básicas cualquiera, existe una tercera actividad cuya producción neta de cada bien es la suma de las producciones netas de dicho bien en aquellas dos actividades.

Geométricamente esto significa que si a y b son actividades básicas posibles, su suma $a + b$, también es una actividad posible (figura 26).

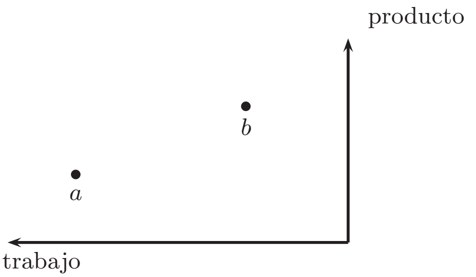


Figura 25. Dos actividades básicas diferentes

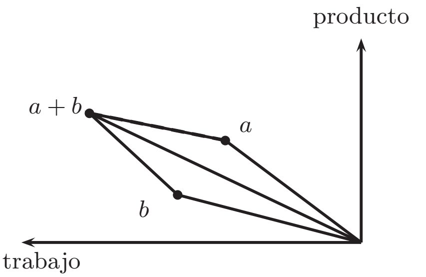


Figura 26. Aditividad de actividades

Obsérvese que esto, a su vez, implica que *no hay interacción entre procesos productivos*: los modos de producción no se afectan unos a otros. En general, aquellos casos en los que existen interacciones (físicas, tecnológicas, etc.) no pueden abarcarse adecuadamente por este modelo. Sin embargo, Koopmans afirma que *en aquellas situaciones en que exista interacción física, la aplicabilidad del presente modelo puede recuperarse en ocasiones reuniendo las actividades interrelacionadas en una única actividad, que tenga como producción neta para cada bien, la suma de las producciones netas de ese bien en cada una de las actividades que la constituyen, una vez tenida en cuenta su interacción.*

Postulado 9 (Proporcionalidad)

Si una actividad (básica o derivada) a es posible, también lo es toda actividad ka para cualquier factor de proporcionalidad k no-negativo. Esto implica, entonces, que la *inactividad* es posible, es decir, que $a = 0$ es posible.

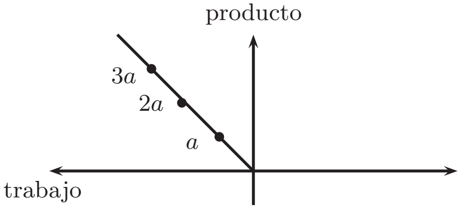


Figura 27. Proporcionalidad de a

En términos geométricos, esto implica que la semirrecta que parte del origen y pasa por a está conformada por actividades posibles (figura 27). En general, este postulado implica lo que se acostumbra a llamar *rendimientos constantes a escala*. Obsérvese que los postulados 7, 8 y 9 implican que los conjuntos de producción son conjuntos de la forma ilustrada en la figura 28.

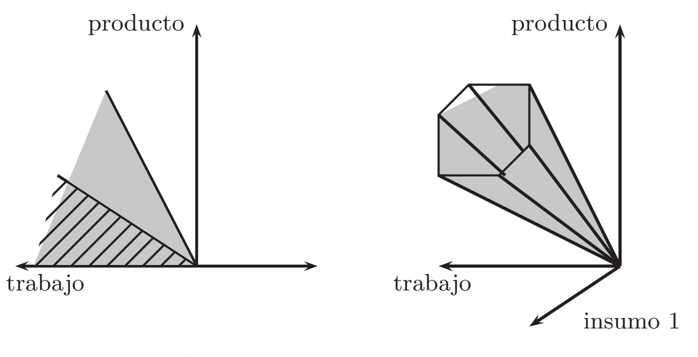


Figura 28. Conjuntos “lineales” de producción

Es conveniente resaltar en este punto que para muchos economistas el término *"linealidad"* se asocia con limitación, restricción e inflexibilidad en las hipótesis. Sin embargo, Koopmans hace la observación de que su "modelo lineal" se relaciona con las hipótesis de proporcionalidad de los insumos y de los productos en cada una de las actividades elementales productivas, y con la hipótesis de que el resultado de llevar a cabo dos o más actividades es la *suma* de los resultados de cada una de estas actividades. Es decir, Koopmans estudia en su modelo *tecnologías con rendimientos constantes a escala (más precisamente, con grado uno de homogeneidad), mas no linealidad en las funciones de producción.* Así, los conos, los poliedros o conjuntos convexos similares son bien asimilados por la teoría de Koopmans, pero también funciones de producción curvilíneas con tales propiedades de homogeneidad.

Postulado 10 (Disponibilidad de recursos)
Cada bien primario puede extraerse de la naturaleza *en cualquier cantidad* no-negativa que exceda una cota superior dada.

Postulado 11 (Sobre la imposibilidad de producción sin utilizar factores)
Ninguna actividad (básica o derivada) es posible si todos sus factores son nulos. Es decir, ningún conjunto de producción puede tener un punto en común con $\mathbb{R}^n_+$ distinto al origen (figura 28).

Postulado 12 (Sobre la posibilidad de producción)
Existe un punto con producción neta no-negativa de todos los bienes deseados, producción positiva de por lo menos uno de ellos, y producción neta igual a cero de los bienes intermedios.

En particular, estos dos últimos postulados establecen que no todos los politopos pueden ser conjuntos de producción admisibles por el modelo.

Definiciones principales e implicaciones del modelo de actividades

Con la escena del modelo de análisis de actividades puesta a través de los doce postulados anteriores estamos listos para enunciar las principales definiciones y proposiciones. Veamos esto.

Definición 7. (Equilibrio de mercado)
Se dice que una combinación de elecciones (una para cada agente) es un *equilibrio de mercado* si la suma neta de todas las cantidades de cada bien, elegidas por los productores y poseedores de recursos, es igual a la de todas las cantidades elegidas por los consumidores.

Definición 8. (Equilibrio competitivo (o walrasiano))
Un *equilibrio competitivo (o walrasiano)* es una combinación de equilibrio de mercado y un sistema de precios (uno para cada bien) tales que si todos los "valores" se calculan a dichos precios,

i) La elección de cada consumidor es preferida o equivalente a cualquier otra elección dentro de su conjunto de consumo.

ii) La elección de cada productor da lugar al máximo beneficio alcanzable dentro de su conjunto de producción.

Es decir, un equilibrio competitivo es una colección de asignaciones de equilibrio de mercado (una para cada agente de la economía), y un sistema de precios para los bienes, tales que cada consumidor esté satisfecho al máximo con su consumo, y cada productor también esté satisfecho, pues su producción le ha maximizado los beneficios de su empresa, dadas las restricciones del mercado.

Ahora: la idea de que, de cierta forma, la competencia perfecta y, por tanto, el equilibrio competitivo, conlleva ciertos objetivos de eficiencia, ha aparecido aquí y allá, tanto en autores clásicos como en los llamados *"neoclásicos"*. Sin embargo, formalizar la noción de eficiencia no fue simple. Pareto sería el primero en introducir la idea de que un equilibrio competitivo permite también comparar los estados de satisfacción de los consumidores, pues notaba que en tal equilibrio no era posible aumentar el nivel de satisfacción de algún consumidor sin que disminuyera el nivel de satisfacción de algún otro. Algunos economistas tales como Allais (1943)[37], Hicks (1939)[38], Kaldor (1939)[39], Lange (1942)[40] y Lerner (1946)[41] encontrarían que, efectivamente, los equilibrios competitivos eran *"óptimos paretianos"*; pero, más que eso, encontraron que todo óptimo paretiano podía también formularse como equilibrio competitivo siempre que las dotaciones de los consumidores fueran redistribuidas convenientemente.

Este último resultado es fundamental, pues no toma como datos institucionales el que la producción la lleven a cabo las empresas, o que los intercambios tengan lugar a través de mercados, sino que parte de los fines a los que deben servir estas u otras instituciones: a la satisfacción de los consumidores. Pero, además, estos dos resultados conducirían directamente a las discusiones que

[37] Allais, M. (1943), *A La Recherche d'une Discipline Économique*, Paris: Imprimerie Nationale.
[38] Hicks, J. (1939), Foundations of Welfare Economics, *Economics Journal*, vol. 49.
[39] Kaldor, N.(1939), Welfare Propositions in Economics, *Economics Journal*, vol. 49.
[40] Lange, O. (1942),The Foundation of Welfare Economics, *Econometrica*, vol. 10, 215-228.
[41] Lerner, A.(1946), *Money*, Encyclopaedia Britannica.

sobre bienestar económico había abierto von Mises (1922, 1932)[42, 43] pues la noción de los precios como señal de información que circula entre los distintos centros de decisión de una economía para asignar eficientemente, estaba en el corazón de la discusión tanto para economías socialistas, como capitalistas, y aún en aquellas en las que algunos mercados son mixtos.

Definición 9. (Óptimo de Pareto)
Un *óptimo de Pareto* (u *óptimo paretiano*) es una combinación de equilibrio de mercado tal que no existe ninguna otra combinación (también de equilibrio de mercado) bajo la cual, si un agente mejora (con respecto a la relación de preferencia), entonces algún otro agente empeora.

Y el siguiente es uno de los dos resultados centrales del modelo de Koopmans:

Teorema 4. ***(Koopmans (1951))***
Si en el modelo (lineal) de análisis de actividades se satisfacen los postulados 1–5, cualquier equilibrio walrasiano es un óptimo de Pareto.

En palabras del propio Koopmans (1951), este teorema

> reduce a sus elementos lógicos esenciales la creencia clásica en la eficiencia de la competencia como mecanismo de asignación de recursos para la producción y el consumo. Establece que, si existe un sistema de precios común que cuando se utiliza para determinar todas las acciones de máximos de beneficios o de utilidad, permite o induce decisiones compatibles (...), dicha utilización de los precios también garantiza el empleo eficiente de los recursos para satisfacer las preferencias de los consumidores, tal como se ha definido.

Sin embargo, es conveniente e importante advertir aquí que un óptimo en el sentido de Pareto puede ser una distribución de riqueza más desigual que lo que alguna noción normativa podría sugerir. Así que el título de "óptimo" puede ser, en este sentido, engañoso. Koopmans recomendó en su época que lo llamasen estado "eficiente desde el punto de vista de la asignación". Pero hoy está tan arraigada la terminología "óptimo de Pareto" que quizás sólo sirva para recordar y homenajear a quien, según la historia oficial, lo introdujera en la historia del pensamiento económico[44].

Finalmente, el segundo y fundamental resultado del modelo de Koopmans es, en cierta forma, un recíproco del teorema anterior:

---

[42] Von Mises, L. (1922), *Socialism: An Economic and Sociological Analysis*, Indianapolis: Liberty Press.
[43] Von Mises, L.(1932), *Epistemological Problems of Economics*, Princeton.
[44] A pesar de que el mismo Léon Walras hubiese anticipado a Pareto en esto (Volumen 3 (Optimización y dinámica), lección 2).

Teorema 5. ***(Koopmans (1951))***
En el modelo (lineal) de análisis de actividades, si se satisfacen los postulados 1–12, entonces a todo óptimo de Pareto se le puede asociar un sistema de precios, no todos nulos, tal que sea un equilibrio walrasiano.

Este teorema es el corazón de lo que hoy se llama "nueva economía del bienestar". En palabras del propio Koopmans, "partiendo del ideal de eficiencia en la asignación, expresada por la noción paretiana de optimalidad, formula criterios más directamente aplicables expresados en términos de precios que implican la realización de aquel ideal bajo ciertas condiciones bien especificadas".

Final

Aunque el modelo de análisis (lineal) de actividades sirve para aproximarnos a una mirada abstracta de las implicaciones sobre los precios de una utilización eficiente de los recursos, y del empleo de precios como medio para sostener usos eficientes, también ha podido utilizarse para darle solución numérica al problema práctico (de programación lineal) de encontrar la asignación más eficiente o beneficiosa al interior de una empresa. En particular, puesto que el modelo insumo-producto de Leontief es sólo un caso especial del análisis (lineal) de actividades, cualquier estudio de aquél podría ser extendido al modelo general. Uno de los propósitos de este programa de investigación ha sido el de proveer de una base empírica que permita estimación numérica. Los conceptos teóricos del análisis de actividades se adaptarían, en principio, a responder tales preguntas de política económica cuantitativa partiendo de variables observables de tipo más o menos agregado.

Debe anotarse, para terminar, que el modelo Arrow-Debreu (1954,1959) que estudiaremos en el Volumen 3 (Optimización y dinámica), es todavía una generalización del modelo de actividades de Koopmans. Aún así, ninguno de estos dos modelos podrá ir más allá de *los rendimientos constantes a escala que son la base misma de la competencia.* Una primera posibilidad de ataque a este problema había sido expuesta, en ciernes, en la teoría de juegos de von Neumann y Morgenstern de 1944. Pero este trabajo no tuvo, inicialmente, el impacto que merecía, por razones que expondremos posteriormente (Volumen 3 (Optimización y dinámica), lección 3).

Ejercicios complementarios

1) ¿Cuáles de los siguientes conjuntos son convexos?:

a) $\{(x, y) \in \mathbb{R}^2 / x = 1\}$

b) $\{(x, y) \in \mathbb{R}^2 / y = x^2\}$

c) $\{(x, y) \in \mathbb{R}^2_+ / y \geq \frac{1}{x}\}$

d) $\{(x, y) \in \mathbb{R}^2_+ / xy + y^2 + 1 \geq 0\}$

e) $\{(x, y) \in \mathbb{R}^2 / x > 1, y > 3\}$

f) $\{(x, y) \in \mathbb{R}^2 / 3x + 2y \geq 7\}$

g) $\{(x, y) \in \mathbb{R}^2 / 3x + 2y = 7\}$

h) $\{(x, y) \in \mathbb{R}^2 / 3x + 2y \leq 7\}$

i) $\{(x, y, z) \in \mathbb{R}^3 / x^2 + y^2 + z^2 \leq 1\}$

j) $\{(x, y, z) \in \mathbb{R}^3 / 35x - 2y + 2z \leq 1\}$

Ilustre con una gráfica en cada caso. Si el conjunto no es convexo, halle su envolvente.

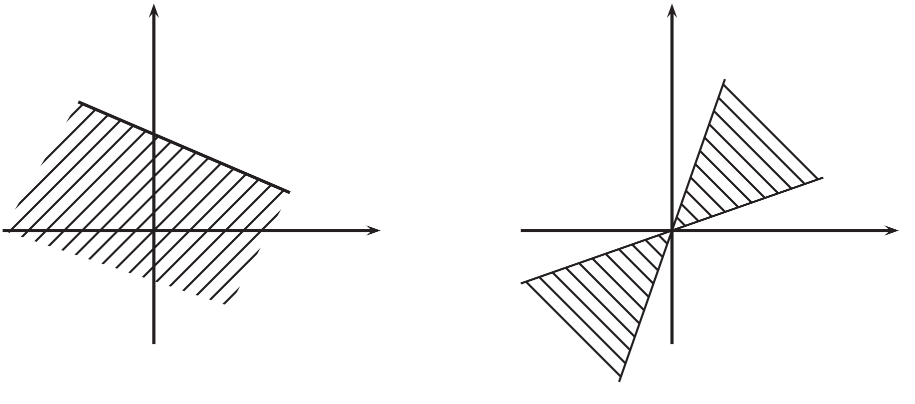
2) ¿Cuáles de los siguientes conjuntos son politopos, poliedros o conos convexos?

a) $\mathbb{R}^n_+$

b) $\{(x, y) \in \mathbb{R}^2 / x \geq 0, y \geq 0\} \cup \{(x, y) \in \mathbb{R}^2 / x \leq 0, y \leq 0\}$

c) $\{(x, y, z) \in \mathbb{R}^3 / 3x + 2y + 5z \leq 3,\ x - y + 2z \geq 5,\ x + 3y - z \leq 1\}$

3) ¿Cuáles de los siguientes gráficos corresponden a politopos, poliedros o conos convexos?



4) Pruebe que la suma y la intersección de dos conos convexos es, de nuevo, un cono convexo.

5) Resuelva los siguientes problemas de programación lineal *mediante re-*

presentación gráfica:

a) $$\begin{aligned} \text{Maximizar} \quad & 3x + 2y \\ \text{sujeta a} \quad & 2x + y \leq 6 \\ & x + 2y \leq 8 \\ & x \geq 0 \\ & y \geq 0 \end{aligned}$$

b) $$\begin{aligned} \text{Maximizar} \quad & 2x + 10y \\ \text{sujeta a} \quad & x + y \leq 3 \\ & x + 2y \leq 5 \\ & x \geq 0 \\ & y \geq 0 \end{aligned}$$

c) $$\begin{aligned} \text{Minimizar} \quad & 3x + 2y \\ \text{sujeta a} \quad & x - y \geq 1 \\ & x \geq 0 \\ & y \geq 0 \end{aligned}$$

d) $$\begin{aligned} \text{Maximizar} \quad & 7x + 2y \\ \text{sujeta a} \quad & 3x + y \geq 5 \\ & x + y \geq 2 \\ & x \geq 0 \\ & y \geq 0 \end{aligned}$$

6) Suponga un individuo que obtiene ganancias de la venta de dos clases de libros: de economía y de literatura. Cada libro de economía tiene un precio de venta de $30,000$ pesos, y cada libro de literatura tiene un precio de venta de $15,000$ pesos. Si el máximo número de libros (de economía y de literatura) que esta persona puede vender al mes es 100, y dado que la producción mensual de libros de economía es el doble de la producción de libros de literatura ¿cuál es la ganancia máxima mensual del vendedor? Plantee y resuelva el problema de programación lineal a partir de la información anterior.

7) El departamento químico de una empresa decide comprar por lo menos $8,000$ probetas pequeñas y $5,000$ grandes. Puede aprovechar un precio especial si adquiere por lo menos $15,000$ piezas. La empresa planea comprar, a lo más, el triple de probetas grandes que de pequeñas. ¿Cuántas probetas de cada tamaño es necesario comprar para satisfacer los requerimientos de la empresa? Dar al menos cinco soluciones enteras. Dibujar la región de posibles soluciones.

8) Suponga una economía descrita mediante el modelo "keynesiano" IS-LM de la siguiente manera:

$$i = \frac{1}{\beta}\left[F - (1-c)Y\right]$$

$$i = \frac{1}{h}\left[kY - \frac{M}{P}\right]$$

donde i = tasa de interés, F = "índice" de política fiscal (gasto), Y = ingreso agregado, M = saldos monetarios nominales, P = nivel de precios, y $\beta > 0$, $h > 0$, $k > 0$, $0 < c < 1$ son constantes. Si $\beta = 0.6$, $h = 0.2$,

$k = 0.3$, $c = 0.8$, $\dfrac{M}{P} = 200$ y $F = 600$, calcule, si existen, los valores correspondientes de i y Y para que esta economía esté en "equilibrio". Represente la solución gráficamente.

9) Supongamos que una empresa que produce automóviles obtiene un beneficio de 100 dólares en cada auto tipo A, 200 dólares en cada auto tipo B, y 400 dólares en los de tipo C. Los autos recorren 20, 17 y 14 millas por galón, respectivamente, pero regulaciones del gobierno insisten en que el promedio debe ser de 18 millas por galón. La planta puede ensamblar un auto tipo A en 1 día, uno tipo B en 2 días y uno tipo C en 3 días. ¿Cuál es el beneficio máximo de ocho meses (de 30 días cada uno) de trabajo? [Indicación: el problema es

$$\begin{array}{ll} \text{Maximizar} & 100x + 200y + 400z \\ \text{sujeta a} & 20x + 17y + 14z \geq 18(\,x + y + z\,) \\ & x + 2y + 3z \leq 240 \\ & x \geq 0 \\ & y \geq 0 \\ & z \geq 0 \end{array}$$

donde x es el número de autos tipo A, y es el número de autos tipo B; y z es el número de autos tipo C].

10) Resuelva los siguientes juegos de matriz:

a) $A = \begin{bmatrix} 4 & 0 \\ 6 & 3 \end{bmatrix}$ b) $A = \begin{bmatrix} 4 & -3 \\ 0 & 2 \end{bmatrix}$

*11) Pruebe que el valor del juego de matriz

$$A = \begin{bmatrix} 6 & -2 & 3 \\ -4 & 5 & 4 \end{bmatrix}$$

es $v = \frac{22}{17}$, y las probabilidades del jugador 1 son $p_1 = \frac{9}{17}$, $p_2 = \frac{8}{17}$. ¿Cuáles son las probabilidades del jugador 2?

12) Construya gráficamente conjuntos de consumo y producción *determinados paramétricamente*, y que satisfagan los postulados del modelo de análisis de actividades de Koopmans.

Bibliografía

ALCHIAN, A. (1950): "Uncertainty, Evolution and Economic Theory", *Journal of Political Economics*, vol. 58(3), 211–221.

ALCHIAN, A. AND H. DEMSETZ (1972): "Production, Information Costs and Economic Organization", *American Economic Review*, vol. 62, 777–95.

ALLAIS, M. (1943): *A La Recherche d'une Discipline Économique*. Paris: Imprimerie Nationale.

——— (1953): "Le Comportement de l´Homme Rationel Devant le Risque: Critique des Postulats et Axioms de l'École Americaine", *Econometrica*, vol. 21.

ARROW, K. (1951a): "Alternative Approaches to the Theory of Choice in Risk-Taking Situation", *Econometrica*, vol. 19, 404–437.

——— (1951b): *An Extension of the Basic Theorems of Classical Welfare Economics*. Berkeley: University of California Press.

ARROW, K., S. CHENERY AND R. SOLOW (1961): "Capital Labor Substitution and Economic Efficiency", *Review of Economic Studies*, vol. 63, 225–250.

ARROW, K. AND G. DEBREU (1954): "Existence of an Equilibrium for a Competitive Economy", *Econometrica*, vol. 22., 265–290.

ARROW, K. AND A. ENTHOVEN (1961): "Quasi-Concave Programming", *Econometrica*, vol. 29, 779–800.

ARROW, K., L. HURWICZ AND H. USAWA (1961): "Constraint Qualifications in Maximization Problems", *Naval Research Logistics Quarterly*.

ARROW, K. AND M. D. INTRILIGATOR (1981-1991): *Handbook of Mathematical Economics*. 4 Vols. North Holland: Amsterdam.

ARTHUR, B. (1989): "Competing Technologies, Increasing Returns, and Lock-in by Historical Events", *The Economic Journal*, vol. 99, 116–131.

ARTHUR, W. B., S. N. DURLAUF AND D. A. LANE (1997): *The Economy as an Evolving Complex System II.* SFI Studies in the Sciences of Complexity, New York: Addison Wesley Longman.

AUMANN, R. J. (1964): "Markets with a Continuum of Traders", *Econometrica*, vol. 32, 39–50.

——— (1976): "Agreeing to Disagree", *Annals of Statistics*, vol. 4(6), 1236–1239.

BANERJEE, A. V. (1992): "A Simple Model of Herd Behavior", *The Quarterly Journal of Economics*, vol. 107(3), 797–817.

BARRO, R. AND X. SALA-I-MARTIN (1995): *Economic Growth.* New York: McGraw-Hill Advance Series in Economics.

BAUMOL, W. AND S. GOLDFELD (1968): *Precursors in Mathematical Economics.* London: School of Economics.

BELLMAN, R. E. (1957): *Dynamic Programming.* Princeton: Princeton University Press.

BENTHAM, J. (1789): *An Introduction to the Principles of Morals and Legislation.* Kitchener: Batoche Books.

BERGE, C. (1959): *Topological Spaces: Including a Treatment of Multi-valued Functions, Vector Spaces and Convexity.* Paris: Dunod.

BERGSON, A. (1938): "A Reformulation of Certain Aspects of Welfare Economics", *Quarterly Journal of Economics*, vol. 52, 310–334.

BERNOULLI, D. (1954): "Specimen Theoriae Novae de Mensura Sortis", *Econometrica*, vol. 22, Trad. L. Sommer versión 1738.

BERTRAND, J. (1883a): "Théorie des richesses: revue de théories mathématiques de la richesse sociale par Léon Walras et recherches sur les principes mathématiques de la théorie des richesses par Augustin Cournot", *Journal des Savants*, vol. 67, 499–508.

——— (1883b): "Théorie Mathématique de la Richesse Sociale", *Journal des Savants*, vol. 67.

BERTSEKAS, D. (1987): *Dynamic Programming.* Englewoods Cliffs: Prentice-Hall, Inc.

——— (2007): *Dynamical Programming and Optimal Control*, vol. I, II. Athena Scientific, third edn.

Böhm-Bawerk, E. (1889): *Capital and Interest*. South Holland: Libertarian Press.

Bickchandani, S., D. Hirshleifer and I. Wech (1992): "A Theory of Fads, Fashion, Custom, and Cultural Exchange in Informational Cascades", *Journal of Political Economy*, vol. 100(5), 992–1026.

Blanchard, O. and S. Fischer (1989): *Lectures on Macroeconomics*. MIT Press.

Blume, L. and S. D. (eds.) (2001): *Social Dynamics*. New York.

Boldrin, M. (1983): *Applying Bifurcation Theory: Some Simple Results on the Keynesian Business Cycle*. Note di Lavoro 8403, University di Venezia.

Borel, m. (1921): "La théorie du Jeux et las Equations Intégrales à Noyau Symétrique", *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences*, vol. 173, 1304–1308.

Bourbaki, N. (1939): *Éléments de Mathématiques*. París: Hermann.

Bowley, A. L. (1924): *The Mathematical Groundwork of Economics*. Oxford: Oxford University Press.

Boyer, C. B. (1959): *History of the Calculus and Its Conceptual Development*. New York: Dover.

Brown, D. and A. Robinson (1972): *A Limit Theorem on the Cores of Large Standard Exchange Economies*. EEUU: Proceedings of the National Academy of Sciences.

Cassel, G. (1918): *Theory of Social Economy*. New York: Harcourt, Brace and Company, 1932 edn.

Cauchy, A. (1821): *Cours d'Analyse Algébrique*. Paris: Chez Debure Frère.

Cayley, A. (1841): "A Theorem in the Geometry of Position", *Cambridge Mathematical Journal*, vol. 2, 267–271.

——— (1858): "A Memoir on the Theory of Matrices", *The Royal Society of London*, vol. 148, 17–37.

Cerdá, E. (2001): *Optimización Dinámica*. Madrid: Pearson Educación, S. A.

Chamberlin, E. (1933): *The Theory of Monopolistic Competition*. Cambridge (Ma.): Harvard University Press, 7 edn.

——— (1956): *The Theory of Monopolistic Competition : a Reorientation of the Theory of Value*. Cambridge (Ma.): Harvard Economic Studies.

Chiang, A. C. (1992): *Elements of Dynamic Optimization*. UK: McGraw-Hill, Inc.

Clark, J. B. (1889): "Possibility of a Scientific Law of Wages", *Publications of the American Economic Association*, vol. 4(1).

——— (1891): "Distribution as Determined by a Law of Rent", *Quarterly Journal of Economics*, vol. 5(3), 289–318.

——— (1899): *The Distribution of Wealth: A Theory of Wages, Interest and Profits*. New York: Macmillan.

Coase, R. H. (1937): "The Nature of the Firm", *Economica*, vol. 4(16), 386–405.

Cobb, C. and P. Douglas (1928): "A Theory of Production", *American Economic Review*, vol. 18, 139–65.

Courant, R. (1937): *Differential and Integral Calculus*, vol. I–II. New York: Wiley.

Cournot, A. (1838): *Researches into the Mathematical Principles of the Theory of Wealth*. London: Macmillan.

Cramer, G. (1750): *Introduction a l'Analyse des Lignes Courbes Algebriques*.

Dana, R. and P. Malgrange (1984): *The Dynamics of a Discrete Version of a Growth Cycle Model*. New York: W.H. Freeman and Company.

Dantzig, G. B. (1949): "Programming in a Linear Structure", *Econometrica*, vol. 17, 73–74.

Davidson, P. (1972): "A Keynesian View of Friedman's Theoretical Framework for Monetary Analysis", *Journal of Political Economy*, vol. 80(5), 864–82.

——— (1994): "The Asimakopulos View of Keynes's General Theory", in *Investment and Employment in Theory and Practice*, ed. by Harcourt, and Roncaglia.

Day, R. and J. Shafer (1985): "Keynesian Chaos", *Journal of Macroeconomics*, vol. 7(2), 277–295.

Day, R. H. (1982): "Irregular Growth Cycles", *American Economic Review*, vol. 72(3), 406–14.

——— (1983): "The Emergency of Chaos from Classical Economic Growth", *Quarterly Journal of Economics*, vol. 98(2), 201–13.

Deaton, A. (1992): *Understanding Consumption*. Oxford: Clarendon Press.

Debreu, G. (1954): "Representation of a Preference Ordering by a Numerical Function", in *Decision Processes*, ed. by R. M. Thrall, C. H. Coombs, and R. L. Davis. New York: Wiley.

——— (1959): *The Theory of Value, An Axiomatic Analysis of Economic Equilibrium*. New Haven and London: Yale University Press.

——— (1962): "New Concepts and Techniques for Equilibrium Analysis", *Economic Review*, vol. 3(3), 257–273.

——— (1974): "Excess Demand Functions", *Journal of Mathematical Economics*, vol. 1, 15–23.

——— (1991): "The Mathematization of Economic Theory", *The American Economic Review*, vol. 81(1), 1–7.

Debreu, G. and T. Koopmans (1982): *Additively Decomposed Quasiconvex Functions*. Mathematical Programming: Scientific Papers of Tjalling C. Koopmans (1985), Springer Verlag.

Descartes, R. (1637): *Discourse on the Method of Rightly Conducting the Reason and Seeking the True in the Sciences with Ap-plications: Dioptrics, Meteorology and Geometry*.

Dierker, E. (1972): "Two Remarks on the Number of Equilibria of an Economy", *Econometrica*, vol. 40(5), 951–53.

Dixit, A. K. (1990): *Optimization in Economic Theory*. New York: Oxford University Press, 2nd. edn.

Dupuit, J. (1844): "On the Measurement of the Utility of Public Works", *International Economic Papers*, vol. 2.

——— (1933): *De l'Utilité et de sa Mesure*. Turin: Marie de Bernardi.

Durlauf, S. N. and W. Brock (2001): "Discrete Choice with Social Interactions", *Review of Economic Studies*, vol. 68(2), 235–60.

Edgeworth, F. (1881): *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*. London: Kegan Paul and Co.

——— (1897): "The Pure Theory of Monopoly", *Giornale degli Economisti*, vol. 40, 13–31.

——— (1899): "Utility", in *Dictionary of Political Economy*, ed. by R. H. I. Palgrave, vol. 3. London: Macmillan.

Edgeworth, F. Y. (1877): *New and Old Methods of Ethics*. London: James Parker.

Euler (1748): *Introductio in Analysin Infinitorum*. Lausanne: M.-M. Bousquet.

Evans, G. C. (1924): "The Dynamics of Monopoly", *American Mathematical Monthly*, vol. 31(2), 77–83.

Fermat, P. (1679): *Ad Locos Planos et Solidos Isagoge*. Toulouse.

——— (1891-1912): *OEuvres de Fermat*. Vol. I-IV. Paris: Gauthier-Villars.

Fishburn, P. C. (1994): *Utility and Subjective Probability*. Elsevier: Amsterdam.

Fisher, F. (1966): *The Identification Problem in Econometrics*. New York: Mc Graw Hill.

Fisher, I. (1926): *Mathematical Investigations in the Theory of Value and Prices*. New Haven: Yale University Press.

——— (1930): *The Theory of Interest: As Determined by the Impatience to Spend Income and Opportunity to Invest it*. New York: Macmillan.

Fréchet, M. (1906): *Sur quelques points du calcul fonctionnel*. Springer Wien.

Friedman, M. (1953): *Essays in Positive Economics*. Chicago: University of Chicago Press.

——— (1973a): *Una Teoría de la Función Consumo*. Traducción de Lorenzo Betancor Curbelo. Madrid: Alianza Editorial.

——— (1973b): *'Windfalls', the 'Horizon', and Related Concepts on the Permanent Income Hypothesis*. Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld. Stanford University Press.

——— (1982): *Teoría de los precios*. Traducción de José Vergara. Alianza Universidad.

Galilei, G. (1891-1912): *Dialogo Sopra i due Massini Sistemi del Mondo*. Firenze.

GAUSS, K. F. (1809): *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium.*

GINTIS, H. (2005): "The Dynamics of General Equilibrium", *Economic Journal*, vol. 117(523), 1280–1309.

GOODWIN, R. (1951): "The Nonlinear Accelerator and the Persistence of Business Cycles", *Econometrica*, vol. 19, 1–17.

GORDON, R. A. (1976): "Rigor and Relevance in a Changing Institutional Setting", *American Economic Review*, vol 66, 1–14.

GOSSEN, H. (1854): *Entwicklung der Gesetze des menschlichen Verkehrs und der daraus fliessenden Regeln für menschliches Handeln.* Braunschweig.

GRANDMONT, J. M. (1985): "On Endogenous Competitive Business Cycles", *Econometrica*, vol. 53, 995–1045.

GRASSMANN, H. (1844): *Die lineare Ausdehnungslehre.* Leipzig: Wiegand.

——— (1861): *Lehrbuch der Mathematik für höhere Lehranstalten.* Berlin: Enslin.

GROBMAN, D. (1959): *Homeomorphisms of Systems of Differential Equations.* Dokl. Akad. Nauk.

GUCKENHEIMER, J. AND P. HOLMES (1983): *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields.* New York: Springer–Verlag.

HADAMARD, J. (1910): *Leçons sur le calcul des variations.* Paris: Éditions Jacques Gabay.

HAHN, F. (1962): "On the Stability of Pure Exchange Equilibrium", *International Economic Review*, vol. 3, 206–213.

——— (1982): "The neo-Ricardians", *Cambridge Journal of Economics*, vol. 6, 353–74.

HAHN, F. H. AND T. NEGISHI (1962): "A Theorem on Non-Tâtonnement Stability", *Econometrica*, vol. 30(3), 463–469.

HALL, H. AND S. KNIGHT (1948): *Álgebra Superior.* México: Uteha.

HALL, R. (1978): "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence", *The Journal of Political Economy*, 86(6), 971–987.

Hall, R. L. and C. J. Hitch (1939): "Price Theory and Business Behavior", *Oxford Economic Papers*, vol. 2, 12–45.

Hamilton, D. P. (1990): "The SSC Takes On a Life of Its Own", *Science*, vol. 249(4970), 731–732.

Hansen, A. (1947): *Keynes on Economic Policy.* en Harris, editor, New Economics.

——— (1953): *A Guide to Keynes.* New York: Mc Graw Hill.

Harrod, R. F. (1932): "Decreasing Costs: An Addendum", *Economics Journal*, vol. 42.

——— (1937): "Mr. Keynes and Traditional Theory", *Econometrica.*

Hart, S. and A. Neyman (1988): "Values of Non-Atomic Vector Measure Games", *Journal of Mathematical Economics*, vol. 17(1), 31–40.

Hartman, P. (1963): "On the Local Linearization of Differential Equations", *Proceedings of the American Mathematical Society*, vol. 14, 568–573.

Hawkins, D. and H. Simon (1949): "Some Conditions of Macroeconomics Stability", *Econometrica*, vol. 17, 245–248.

Heath, T. L. (1908): *The Thirteen Books of Euclid.* Cambridge: The University Press.

Herstein, I. and Milnor (1953): "An Axiomatic Approach to Measurable Utility", *Econometrica*, vol. 21, 291–297.

Hicks, J. and R. Allen (1934): "A Reconsideration of the Theory of Value", *Economica*, vol. 1(1), 52–76.

Hicks, J. R. (1932): "Marginal Productivity and the Principle of Variation", *Economica*, vol. 12, 79–88.

——— (1936): "Mr. Keynes's Theory of Employment", *Economic Journal*, vol. 46, 238–253.

——— (1937): "Mr. Keynes and the Classics: A Suggested Interpretation", *Econometrica*, vol. 5(2), 147–159.

——— (1939): "Foundations of Welfare Economics", *Economics Journal*, vol. 49.

——— (1941): "The Rehabilitation of Consumer's Surplus", *Review of Economic Studies*, vol. 8(2), 108–116.

——— (1946): "The Generalized Theory of Consumer's Surplus", *Review of Economic Studies*, vol. 13(2), 68–74.

——— (1950): *A Contribution to the Theory of the Trade Cycle*. London: Clarendon Press.

——— (1976): *"Revolutions in Economics" in Method and Appraisal in Economics*. Cambridge, (Ma.): Cambridge University Press.

——— (1980): "IS-LM: An Explanation", *Journal of Poskeynesian Economics*, vol. 3, 139–155.

——— (1988): "Towards a More General Theory", in *Finance Constraints, Expectations and Macroeconomics*, ed. by Kohn, and Tsiang.

Hilbert, D. (1899): *Fundamentos de la Geometría*. Leipzig: B.G. Teubner.

Hirsch, M. and S. Smale (1974): *Differential Equations, Dynamical Systems, and Linear Algebra*. New York: Academic Press.

Hobson, J. (1891): "The Law of the Three Rents", *Quarterly Journal of Economics*, vol. 5.

Hoffman, K. and K. Ray (1971): *Linear Algebra*. Prentice Hall.

Hotelling, H. (1929): "Stability in Competition", *Economic Journal*, vol. 39(153), 41–57.

——— (1932): "Edgeworth's Taxation Paradox and the Nature of Supply and Demand Functions", *Journal of Political Economy*, vol. 60, 577–616.

Intriligator, M. D. (1971): *Mathematical Optimization and Economic Theory*. Englewood Cliffs, New Jersey: Prentice-Hall.

Jaffé, W. (1973): "Léon Walras's Role in the 'Marginal Revolution' of the 1870's", in *The Marginal Revolution in Economics:*, ed. by R. D. Collison Black, A. W. Coats, and C. Goodwin. Durham, NC: Duke University Press.

Jensen, M. C. and W. H. Meckling (1976): "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure", *Journal of Financial Economics*, vol. 3.

Jevons, W. S. (1871): *The Theory of Political Economy*. New York: A.M. Kelley.

Kahneman, D. and D. Lovallo (1993): "Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking", *Management Science*, vol. 39.

Kahneman, D. and A. Tversky (1979): "Prospect Theory: Analysis of Decision under Risk", *Econometrica*, vol. 47.

Kakutani, S. (1941): "A Generalization of Brouwer's Fixed Point Theorem", *Duke Mathematical Journal*, vol. 8(3), 457–459.

Kaldor, N. (1939): "Welfare Propositions in Economics", *Economics Journal*, vol 49.

——— (1940): "A Model of the Trade Cycle", *The Economic Journal*, vol. 50, 78–92.

——— (1966): *Causes of the Slow Rate of Economic Growth in the UK*. UK: Cambridge.

Kandori, M., G. Mailath and R. Rob (1991): "Learning, Mutation, and Long Run Equilibria in Games", *Econometrica*, vol. 61(1), 29–56.

Kantorovich, L. V. (1939): "The Mathematical Method of Production Planning and Organization", *Management Science*, Vol. 6(4), 363–422.

Kaplan, W. (1991): *Advanced Calculus*. Addison Wesley, 4th edn.

Karush, W. (1939): *Minima of Functions of Several Variables with Inequalities as Side Constraints*. M. Sc. Dissertation, Department of Mathematics, University of Chicago.

Kehoe, T. J. (1985): "Multiplicity of Equilibria and Comparative Statics", *Quarterly Journal of Economics*, vol. 100(1), 119–147.

Kepler, J. (1619): *Harmonies of the World*. American Philosophical Society (1997).

Keynes, J. (1932): "The World's Economic Outlook", *Atlantic Monthly*, May.

——— (1936): *The General Theory of Employment, Interest and Money*. London: Macmillan.

——— (1981): *Teoría general de la ocupación, el interés y el dinero*. FCE, México: Traducción de Eduardo Hornedo.

Kolgomorov, A. N. (1973): *La matemática : su contenido, métodos y significado*. España: Alianza.

Koopmans, T. C. (1951): *Activity Analysis of Production and Allocation*. New York: Wiley & Sons.

——— (1953): "Activity Analysis and its Applications", *American Economic Review*, vol. 43(2), 406–414.

——— (1957): *Three Essays on the State of Economic Science*. New York: McGraw-Hill.

Kreps, D. and R. Wilson (1982): "Reputation and Imperfect Information", *Journal of Economic Theory*, vol. 27(2), 253–279.

Kuhn, H. W. and A. W. Tucker (1951): *Linear Inequalities and Related Systems*. Princeton, New Jersey: Princeton University Press.

——— (1956): *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability: Nonlinear Programming*. Berkeley: University of California Press.

Kuratowski, K. (1921): "Sur la Notion de L'Ordre dans la Théorie des Ensembles", *Fundamenta Mathematicae*, vol. 2, 161–171.

Lagrange (1788): *Mécanique Analytique*. París. Blanchard.

Lange, O. (1936): "On the Economic Theory of Socialism", *Review of Economic Studies*, vol. 4(1), 53–71.

——— (1942a): "The Foundation of Welfare Economics", *Econometrica*, vol. 10, 215–228.

——— (1942b): "The Stability of Economic Equilibrium", *Econometrica*, vol. 10, 176–177.

Laplace, P. S. (1772): *Mécanique Celeste*. Paris: Bachelier.

——— (1799): *Mécanique Celeste*. Paris: Bachelier.

Larew, G. A. (1919): "Necessary Conditions in the Problems of Mayer in the Calculus of Variations", *Transactions of the American Mathematical Society*, vol. 20.

Leonid, V. K. (1975): *Las Matemáticas en la Economía: Logros, Dificultades, Perspectivas.*

Leontief, W. (1933): "The Use of Indifference Curves in the Analysis of Foreign Trade", *Quarterly Journal of Economics*, vol. 97, 493–503.

——— (1936a): "Composite Commodities and the Problem of the Index Numbers", *Econometrica*, vol. 4(1), 39–59.

——— (1936b): "Quantitative Input and Output Relations in the Economic System of the United States", *The Review of Economics and Statistics*, vol. 18(3), 105–125.

——— (1946): "The Pure Theory of the Guaranteed Annual Wage Contract", *Journal of Political Economy*, vol. 54, 99–150.

——— (1971): "Theoretical Assumptions and Non-observed Facts", *American Economic Review*, vol. 61(1), 1–7.

——— (1993): *Análisis Ecónomico Input-Output*. Madrid: Editorial Planeta.

Lerner, A. (1932): "The Diagrammatical Representation of the Cost Conditions in International Trade", *Economica*, vol. 12, 346–56.

——— (1933): "The Diagrammatical Representation of Elasticity of Demand", *Review of Economic Studies*, vol. 1(33), 39–44.

——— (1937): "Statics and Dynamics in Socialist Economics", *Economic Journal*, vol. 4, 72–76.

——— (1938): "Theory and Practice of Socialist Economics", *Review of Economic Studies*, vol. 6(1), 71–5.

——— (1944): "Interest Theory: Supply and Demand for Loans or Supply and Demand for Cash?", *Review of Economic Studies*, vol. 26, 88–91.

——— (1946): "Money", Encyclopaedia Britannica.

——— (1951): *The Economics of Employment*. New York: Mc Graw Hill.

——— (1952a): "The Essential Properties of Interest and Money", *Quarterly Journal of Economics*, vol. 66, 172–93.

——— (1952b): "Factor Prices and International Trade", *Economica*, vol. 19, 67–84.

L'Hospital, G. (1696): *Analyse des Infiniment Petits pour L' Intelligence des Courbes*. Paris: Montalant.

Li, T. and J. Yorke (1975): "Period Three implies Chaos", *American Mathematical Monthly*, vol. 82, 985–992.

Londoño, S. (1959): *Matemática Moderna. Aritmética y Geometría*. Medellín: Bedout.

Lotka, A. (1925): *Elements of Physical Biology*. Baltimore: Williams and Wilkins.

Luce, R. and H. Raiffa (1957): *Games and Decisions*, vol. I–II. New York: Wiley.

Machlup, F. (1935): “The Commonsense of the Elasticity of Substitution”, *Review of Economic Studies*, vol. 2, 202–13.

——— (1946): “Marginal Analysis and Empirical Research”, *American Economic Review*, vol. 36.

Malthus, T. (1798): *Essay on the Principle of Population*. London: J. Johnson.

Mandelbrot, B. (1983): *The Fractal Geometry of Nature*. New York: W.H. Freeman and Company.

Mangasarian, O. (1966): “Sufficient Conditions for the Optimal Control of Nonlinear Systems”, *Siam Journal of Control*, vol. 4.

Mantel, R. (1974): “On the Characterization of Aggregate Excess Demand”, *Journal of Economic Theory*, vol. 7.

Marshall, A. (1890): *Principles of Economics*. London: Macmillan and Co.

Marx, K. (1984): *El Capital. Crítica de la Economía Política*. FCE, México: Traducción de Wenceslao Roces. Vol. 1.

Mas-Colell, A. (1975): “On the Equilibrium Price Set of an Exchange Economy”, *Journal of Mathematical Economics*.

——— (1985): *The Theory of General Economic Equilibrium: A Differentiable Approach*. Cambridge University Press.

Maxwell, J. C. (1865): *A Dynamical Theory of the Electromagnetic Field*. London: The Society.

Maynard Smith, J. and G. R. Price (1973): “The Logic of Animal Conflict”, *Nature*, vol. 246(5427), 15–18.

McFadden, D. (1963): “Constant Elasticity of Substitution Production Functions”, *Review of Economic Studies*, vol. 30.

Mckenzie, L. W. (1954): “On Equilibrium of Graham’s Model of World Trade and Other Competitive Systems”, *Econometrica*, vol. 22, 147–161.

Meade, J. (1937): “A Simplified Model of Mr. Keynes System”, *Review of Economic Studies*, vol. 4, 98–16.

Mill, J. (1848): *The Principles of Political Economy: with some of their applications to Social Philosophy.* London: Longmans, Green and Co., 7 (1909) edn.

Modigliani, F. (1978): "Life Cycle, Individual Thrift, and the Wealth of Nations", *The American Economic Review*, vol. 76(3), 297–313.

Modigliani, F. and A. Ando (1963): "The "Life-Cycle" Hypothesis of Saving: Aggregate Implications and Tests", *The American Economic Review*, vol. 53(1), 55–84.

Modigliani, F. and R. Brumberg (1954): *Utility analysis and the consumption function: an interpretation of cross-section data.* Post-Keynesian Economics. Rutgers University Press.

Monsalve, S. (1999): *Introducción a los Conceptos de Equilibrio en Economía.* Bogotá–Colombia: Universidad Nacional de Colombia, Facultad de Ciencias Económicas.

Monsalve, S. and J. Arévalo (2005): *Un Curso de Teoría de Juegos Clásica.* Bogotá–Colombia: Universidad Externado de Colombia.

Morris, K. (1980): *Mathematics, the Loss of Certainty.* Oxford: University Press.

Muñoz, J. M. (2002): *Introducción a la Teoría de Conjuntos.* Bogotá D.C.: Universidad Nacional de Colombia.

Murray, J. D. (2002): *Mathematical Biology: I. An Introduction.* New York: Springer.

Nash, J. F. (1950): "Equilibrium Points in n-Person Games", *Proceedings of the National Academy of Sciences*, vol. 36(1), 48–49.

Neisser, H. (1932): *Lohnhöhe und Beschäftigungsgrad im Marktgleichgewicht.* Weltwirtschaftliches Archiv.

Nelson, R. and S. Winter (1982): "The Schumpeterian Tradeoff Revisited", *American Economic Review*, vol. 72, 114–32.

Nikaido, H. (1968): *Convex Structures and Economic Theory.* New York: Academic Press.

Ok, E. (2008): *Real Analysis and Probability with Economic Applications.* lecture note in print.

Osborne, M. and A. Rubinstein (1994): *A Course in Game Theory*. Cambridge: MIT Press.

Pantaleoni, M. (1989): *Pure Economics*. Traducción de 1898, London: Macmillan.

Pareto, V. (1906): *Manual of Political Economy*. Traducción de la edición de 1927, New York: Augustus M. Kelley.

Pasinetti, L. (1974): *Growth and Income Distribution: Essays in Economic Theory*. Cambridge: Cambridge University Press.

——— (1984): "El Trabajo Teórico de Sraffa", in *Lecciones de Teoría de la Producción*. México: Fondo de Cultura Económica.

Patinkin, D. (1956): *Money, Interest and Prices*. Evanston Ill: Row, Peterson.

Peano, G. (1959): *Calcolo Geometrico Secondo l'Ausdehungslehre di H. Grassmann, preceduto dalle Operazioni della Logica Deduttiva*. Turin: Bocca.

Perron, O. (1929): "Über Stabilität und asymptotisches Verhalten der Integrale von Differential gleichungs systemen", *Mathematische Zeitschrift*.

Phillips, A. (1958): "The Relationship between Unemployment and the Rate of Change of Money Wages in the United Kingdom 1861-1957", *Economica*, vol. 25(100), 283–299.

Pielou, E. C. (1977): *Introduction to Mathematical Ecology*. New York: Wiley.

Pigou, A. C. (1920): *The Economics of Welfare*. London: Macmillan.

——— (1933): *Theory of Employment*. London: Macmillan and Co.

——— (1941): *Employment and Equilibrium*. London: Macmillan and Co.

Pohjola, M. (1981): "Stable, Cyclic and Chaotic Growth: The Dynamics of a Discrete-Time Version of Goodwin's Growth Cycle Model", *Zeitschrift für Nationalökonomie*.

Poincaré, H. (1952): *Science and Method*. New York: Dover Publications.

Pontryagin, L., V. Boltyanskii, R. Gamkrelidze and E. Mishchenko (1961): *The Mathematical Theory of Optimal Processes*. Gordon and Breach (Traducción de 1986).

Radner, R. (1968): "Competitive Equilibrium under Uncertainty", *Econometrica*, vol. 36(1), 31–58.

Reunión de Profesores, e. (1958): *Curso de Geometría*. París: Leigel.

Ricardo, D. (1817): *Principles of Political Economy and Taxation*. London: John Murray,.

Robbins, L. C. (1930): "On a Certain Ambiguity in the Conception of Stationary Equilibrium", *Economic Journal*, vol. 40, 194–214.

——— (1932): *An Essay on the Nature and Significance of Economic Science*. New York: New York University Press, 1984 edn.

——— (1934): "Remarks Upon Certain Aspects of the Theory of Costs", *Economic Journal*, vol. 44.

Robinson, J. (1933): "Your Position Is Thoroughly Orthodox and Entirely Wrong", *Journal of the History of Economic Thought*, vol. 20(4).

——— (1973): *Collected Economic Papers, Volume IV*. Oxford: Basil Blackwell.

——— (1978): "Keynes and Ricardo", *Journal of Poskeynesian Economics*, vol. 36(1), 31–58.

——— (1979): *The Generalization of the General Theory and Other Essays*. London: Macmillan.

Rothbard, M. (1987): "Time Preference", in *The New Palgrave: A Dictionary in Economics*, ed. by M. M. J. Eatwell, and P. Newman. United States: Macmillan Press Limited.

Rudin, W. (1976): *Principles of Mathematical Analysis*. International Series in Pure and Applied Mathematics.

Russell, B. (1903): *Principles of Mathematics*. Cambridge: University press.

——— (1908): "Mathematical Logic as Based on the Theory of Types", *American Journal of Mathematic*, vol. 30, 222–262.

Samuelson, L. (1998): *Evolutionary Games And Equilibrium Selection*. Massachusetts: MIT Press.

Samuelson, P. (1939): "Interactions between the Multiplier Analysis and the Principle of Acceleration", *The Review of Economic Statistics*, vol. 21, 75–8.

——— (1947): *Foundations of Economic Analysis*. Cambridge: Harvard University Press.

——— (1948): "International Trade and the Equalisation of Factor Prices", *Economic Journal*, vol. 58(230), 163–184.

——— (1958): "An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money", *Journal of Political Economy*, vol. 66(6), 467–482.

Samuelson, P., R. Solow and R. Dorfman (1958): *Linear Programming and Economic Analysis*. The Rand Series.

Sargent, T. J. (1987a): *Dynamic Macroeconomic Theory*. Cambridge: Harvard University Press.

——— (1987b): *Macroeconomic Theory*. London: Academic Press, Inc., second edn.

Savage, L. J. (1954): *The Foundations of Statistics*. New York: Wiley and Sons.

Scarf, H. (1960): "Some Examples of Global Instability of the Competitive Equilibrium", *International Economic Review*, vol. 1, 157–172.

Schlesinger, K. (1933): "On the Production Equations of Economic Value Theory", in *Precursors in Mathematical Economics*. London: School of Economics.

——— (1933-4): *Über die Produktionsgleichungen der Ökonomischen Wertlehre*. Ergebnisse eines Mathematischen Kolloquiums.

Schumpeter, J. (1947): "Theoretical Problems of Economic Growth", *Journal of Economic History*, vol. 7, 1–9.

——— (1971): *Diez Grandes Economistas: de Marx a Keynes*. Traducción de Ángel de Lucas. Alianza Universidad.

Scitovsky, T. (1940): "A Study of Interest and Capital", *Economica*, vol. 7, 293–317.

——— (1942): "A Reconsideration of the Theory of Tariffs", *Review of Economic Studies*, vol. 9, 89–110.

Senior, N. (1836): *An Outline of the Science of Political Economy*. London: Richard Griffin and Company.

SHEPHARD, R. (1953): *Cost and Production Functions*. Princeton: Princeton University Press.

SHUBIK, M. J. (1959): "Edgeworth Market Games", in *Contribution to the Theory of Games IV*, ed. by L. y Tucker. Princeton: Princeton University Press.

SMITH, A. (1776a): *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: W. Strahan.

——— (1776b): *Investigación sobre la naturaleza y causas de la riqueza de las naciones*. México: Fondo de Cultura Económica.

SMITH, E., S. MEYER AND J. HOWARD (1959): *Analytic Geometry*. New York: J. Wiley (1943).

SONNENSCHEIN, H. (1972): "Market Excess Demand Functions", *Econometrica*, vol. 40(3), 549–63.

SPIVAK, M. (1978): *Calculus*. Barcelona: Editorial Reverté.

SRAFFA, P. (1926): "The Laws of Returns under Competitive Conditions", *Economic Journal*, vol. 36, 535–550.

——— (1960): *Production of Commodities by Means of Commodities: Prelude to a Critique of Economic Theory.* Cambridge: Cambridge University Press.

STIGLER, G. (1947): "Professor Lester and the Marginalists", *American Economic Review*, vol. 37, 154 –157.

STIGLER, J. (1954): "The Early History of Empirical Studies of Consumer Behavior", *The Journal of Political Economy*, LXII(2), 95–113.

STOKEY, N. AND R. LUCAS (1993): *Recursive Methods in Economic Dynamics*. Cambridge: Harvard University Press.

TAKAYAMA, A. (1993): *Analytical Methods in Economics*. The University of Michigan Press.

TAKEUCHI, Y. (1976): *Sucesiones y Series*. México: Limusa.

TOBIN, J. (1985): "Neoclassical Theory in America: J.B. Clark and Fisher", *The American Economic Review*, vol. 75(6), 28–38.

——— (1987): *Irving Fisher*. The New Palgrave: A Dictionary in Economics, The Macmillan Press Limited.

TVERSKY, A. AND D. KAHNEMAN (1986): "Rational Choice and Framing of Decisions", *Journal of Business*, vol. 59.

——— (1991): "Loss Aversion in Riskless Choice: A Reference-Dependent Model", *Quarterly Journal of Economics*, vol. 106.

——— (1992): "Advances in Prospect Theory: Cumulative Representation under Uncertainty", *Journal of Risk and Uncertainty*, vol. 5.

UZAWA, H. (1962a): "On the Stability of Edgeworth's Barter Process", *International Economic Review*, vol. 3, 218–32.

——— (1962b): "Walras Existence Theorem and Brouwer's Fixed Point Theorem", *Economics Studies Quarterly*, vol. 8, 59–62.

——— (1963): "On a Two-Sector Model of Economic Growth, II", *Review of Economic Studies*, vol. 30, 105–118.

VANDERMONDE, A. T. (1771): *Memoire sur la Resolution des Equation.* Paris: History Academy Royal Science.

VOLTERRA, V. (1888): *Leçons sur les équations intégrales et les équations intégro-différentielles.* Gauthier-Villars (1913).

——— (1926): "Variazioni e fluttuazioni del numero d'individui in specie animali conviventi", *Mem. R. Accad. Naz. dei Lincei*, vol. 2, 31–113.

VON BAWERK, E. (1889): *Capital and Interest: Positive Theory of Capital.* Traducción de 1959, South Holland, vol. I(11): Libertarian Press.

——— (1984): "The Ultimate Standard of Value", *Annals of the American Academy*, vol. 5, 149–208.

VON HABERLER, G. (1930): "Die Theorie der Komparativen Kosten und ihre Auswertung für die Begründung des Freihandels", *Welfwirtschaftliches Archiv*, vol. 32.

——— (1933): *The Theory of International Trade: with Applications to Comercial Policy.* traducción de 1934, New York: Macmillan.

VON HAYEK, F. (1945): "The Use of Knowledge in Society", *American Economic Review*, vol. 35(4), 519–30.

VON MISES, L. (1922): "Socialism: An Economic and Sociological Analysis", Indianapolis: Liberty Press.

——— (1932): *Epistemological Problems of Economics.* Princeton.

Von Neumann, J. (1928): "Zur Theorie der Gessellshaftspiele", *Mathematische Annalen*, vol. 100, 295–320.

——— (1932): "A Model of General Economic Equilibrium", *Review of Economic Studies, Traducción de 1945*, vol. 13, 1–9.

Von Neumann, J. and O. Morgenstern (1944): *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.

Von Stackelberg, H. (1933): *Zwei Kritische Bemerkungen zur Preistheorie Gustav Cassels*. Zeitschrift für Nationalökonomie.

——— (1934): *Marktform und Gleichgewicht*. Vienna: Julius Springer.

Von Thünen, J. H. (1826): *Der Isolierte Staat in Beziehung auf Landwirthschaft un National Ökonomie*, vol. 1. Traducido al inglés por Carla M. Wartenberg, Pergamon Press Oxford: New York.

Von Wieser, F. (1876): *Über das Verhältnis des Kosten zum Wert ("On the Relation to Cost to Value")*. reimpreso en Wieser, Gesammelte Abhandlungen.

——— (1884): *Über den Ursprung und die Hauptgesetze des wirtschaftlichen Werthes*.

——— (1889): *Natural Value (The Principle of Solution. The Productive Contribution)*. reimpresión de la traducción de 1893, New York: Augustus M. Kelley.

Wald, A. (1934): "Über die Produktionsgleichungen der Ökonomischen Wetlehre I", *Ergebnisse eines Mathematischen Kolloquiums*, vol. 6, 12–20.

——— (1935): "Über die Eindeutige Positive Losbarkeit der Neuen Produktionsgleichungen", *Ergebnisse eines Mathematischen Kolloquiums*, vol. 6, 12–18.

——— (1936): "Über die Produktionsgleichungen der Ökonomischen Wetlehre II", *Ergebnisse eines Mathematischen Kolloquiums*, vol. 7, 1–6.

——— (1942): "Über einige Gleichungssysteme der Mathematischen Ökonomie", *Zeitschrift für Nationalökonomie*, vol. 7, 637–670.

——— (1951): "On Some Systems of Equations of Mathematical Economics", *Econometrica*, vol. 19, 368–403.

Walras, L. (1874a): *Elements of Pure Economics: On the Theory of Social Wealth*. Traducción de la edición de 1926, Homewood, vol. I(11): Richard Irwin.

——— (1874b): *Élements D'economie Politique Pure (Theorie de la Richesse Sociale).* Traducción al español de Julio Segura, 1987, Alianza Editorial.

Weintraub, S. (1958): *An Approach to the Theory of Income Distribution.* Philadelphia: Chilton.

——— (1959): *A General Theory of the Price Level, Output and Income Distribution.* Philadelphia: Chilton.

——— (1961): *Classical Keynesianism: A Plea for its Abandonnment.* Philadelphia: Chilton.

——— (1966): *A Keynesian Theory of Employment, Growth and Income Distribution.* Philadelphia: Chilton.

Whiteman, C. H. (1987): *Problems in Macroeconomic Theory.* London: Academic Press, Inc.

Wicksell, K. (1893): *Value, Capital and Rent.* London: Allen and Unwin. S. Frowein (trad.) en 1954.

——— (1898): *Interest and Prices.* New York: Royal Economic Society.

——— (1900): *Marginal Productivity as the Basis for Economic Distribution.* New York: M.E. Sharpe.

——— (1901a): *Lectures on Political Economy.* London: Routledge & Kegan Paul.

——— (1901b): *On the Problem of Distribution.* London: Routledge & Kegan Paul.

Wicksteed, P. (1894): *An Essay of the Co-ordination of the Laws of Distribution.* London: London School of Economics.

——— (1910): *The Common Sense of Political Economy: including a study of the human basis of economic law.* London: Macmillan and Co.

——— (1914): "The Scope and Method of Political Economy in the Light of the "Marginal" Theory of Value and Distribution", *Economic Journal*, vol. 24(1).

Williamson, O. E. (1986): *The Economic Institutions of Capitalism.* New York: The Free Press.

——— (1991): "Strategizing, Economizing, and Economic Organization", *Strategic Management Journal*, vol. 12, 75–94.

Young, A. (1928): "Increasing Returns and Economic Progress", *Economic Journal*, vol. 38(152), 527–542.

Zahl, S. (1964): "A Deformation Method for Quadratic Programming", *Journal of the Royal Statistical Society*, vol. 26(1), 141–160.

Respuestas

Lección 1. Eliminación gaussiana

Ejercicios 1.

1. a) $a_{11} = 1, a_{12} = 2, a_{13} = 3, a_{1,4} = 1; a_{21} = 2, a_{22} = 1, a_{23} = 2, a_{24} = 5;$ $a_{31} = 1, a_{32} = 3, a_{33} = 1, a_{34} = 2; a_{4,1} = 1, a_{42} = 1, a_{43} = 1, a_{44} = 1.$

Ejercicios 2.

1. a) $x = \frac{3}{2}$, $y = -\frac{2}{3}$.   b) $x = \frac{3}{2}$, $y = -\frac{9}{3}$.

2. a) $x = -4,\ y = -5,\ z = -6$
   b) No tiene solución.
   d) $x = -7,\ y = 6,\ z = 2.$
   e) $\frac{61}{22}t - \frac{3}{2},\quad y = -\frac{16}{11}t,\quad z = t$ para $t \in \mathbb{R}$

3. a) $k \neq 1$ b) $k = 1$ c) Este caso no tiene solución.

4. El número original es 253.

5. Invierte 4 millones y 2 millones, respectivamente.

Ejercicios complementarios

3. c) $x = -5, y = -5;\quad x = \frac{5}{3},\quad y = -\frac{5}{3};\quad x = 0,\quad y = -\frac{5}{3}.$

4. a) $x = t - 2,\ y = t - 5,\ z = t \in \mathbb{R}.$
   d) $x = 2, y = 3,\ z = 5,\ u = -4.$

5. a) $a = 1, b = 7$

b) $a = -1$

c) $a \neq 1$; solución: $x = \frac{b-7}{a+1}$, $y = \frac{7a-14+b-2ab}{a+1}$, $z = \frac{a(7-b)}{a+1}$.

7. El área del primer campo es 1,200 yardas cuadradas, y la del segundo es 600 yardas cuadradas.

9. Hay 48 paquetes de media libra y 72 paquetes de un tercio de libra.

11. El granjero debe destinar 300 m^2 al cultivo de maíz y 200 m^2 al cultivo de trigo.

12. No; por ejemplo, el sistema $x+y=5$, $2x+2y=1$, tiene dos ecuaciones y dos incógnitas, y, sin embargo, no tiene solución.

13. $P_1 = \frac{111}{10}$, $P_2 = \frac{153}{10}$.

Lección 2. Matrices y determinantes

Ejercicios 2.

1. a) Es diagonal, triangular superior, y triangular inferior.

d) Es idéntica, diagonal, triangular superior, y triangular inferior.

i) Es triangular inferior.

l) Es triangular inferior.

Ejercicios 3.

1. a) $\begin{bmatrix} -3 & 7 & 4 \\ 3 & 5 & \frac{1}{2} \\ \frac{13}{3} & 1 & 5 \end{bmatrix}$

b) $\begin{bmatrix} 5 & -3 & 2 \\ 7 & -1 & -\frac{1}{2} \\ \frac{11}{3} & 1 & -11 \end{bmatrix}$

c) $\begin{bmatrix} 5 & -3 & 2 \\ 7 & -1 & -\frac{1}{2} \\ \frac{11}{3} & 1 & -11 \end{bmatrix}$

d) $\begin{bmatrix} 5 & 0 & 14 \\ 22 & 8 & -1 \\ 13 & 6 & -9 \end{bmatrix}$

g) $\begin{bmatrix} 2 & 11 & -5 \\ -2 & -3 & \frac{3}{2} \\ \frac{17}{3} & -4 & -14 \end{bmatrix}$

h) $\begin{bmatrix} -\frac{7}{6} & \frac{15}{2} & \frac{10}{3} \\ \frac{25}{6} & \frac{23}{6} & \frac{7}{12} \\ \frac{35}{6} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$

2. a) $\begin{bmatrix} 11 & 5 \\ 12 & 12 \end{bmatrix}$

b) $\begin{bmatrix} 5 & 6 \\ 3 & 18 \end{bmatrix}$

f) $\begin{bmatrix} 93 & 57 & 63 \\ 108 & 36 & 108 \end{bmatrix}$

3. a) $\begin{bmatrix} 88 & 61 & 176 \\ 30 & 24 & 137 \end{bmatrix}$

b) $\begin{bmatrix} -\frac{2361}{70} & -\frac{49}{12} & 28 & \frac{869}{12} \\ -\frac{2614}{35} & -\frac{769}{12} & -\frac{221}{3} & -\frac{109}{6} \end{bmatrix}$

c) No es posible pues $4AD$ y $3E$ tienen distinto tamaño.

d) $\begin{bmatrix} 372 & 144 & 584 \\ 200 & 96 & 288 \end{bmatrix}$

Ejercicios 4.

1. b) $A^T = \begin{bmatrix} 4 & 5 & -2 & 1 \\ 5 & 8 & 3 & 4 \\ -2 & 3 & 6 & -6 \\ 1 & 4 & -6 & -3 \end{bmatrix}$; es simétrica.

c) $A^T = \begin{bmatrix} 1 & -10 \\ 2 & -12 \\ 3 & 4 \\ 5 & 8 \\ 7 & 14 \end{bmatrix}$; no es simétrica.

f) $A^T = \begin{bmatrix} 5 & 4 & 3 & 2 & 7 \\ 4 & 1 & -1 & 6 & 11 \\ 3 & -1 & 8 & 5 & 1 \\ 2 & 6 & 5 & -9 & 6 \\ 7 & 11 & 1 & 3 & 10 \end{bmatrix}$; no es simétrica.

2. a)
   - $ABC = \begin{bmatrix} 151 & 138 \\ 291 & 276 \end{bmatrix}$
   - $CB^T A^T = \begin{bmatrix} 151 & 291 \\ 138 & 276 \end{bmatrix}$
   - $BCA = \begin{bmatrix} 197 & 26 & 192 \\ 154 & 22 & 154 \\ 207 & 30 & 208 \end{bmatrix}$
   - $C^T B^T A^T = \begin{bmatrix} 151 & 291 \\ 138 & 276 \end{bmatrix}$

3. $Tr(A) = 5, \;\; Tr(B) = 0, \;\; Tr(C) = 12$

4. $AB = \begin{bmatrix} \begin{bmatrix} -8 & 11 \\ 9 & 4 \end{bmatrix} & \begin{bmatrix} 6 & 4 \\ -15 & 21 \end{bmatrix} \\ \begin{bmatrix} -2 & -29 \\ -2 & -1 \end{bmatrix} & \begin{bmatrix} 5 & -17 \\ 1 & 7 \end{bmatrix} \end{bmatrix}$

Ejercicios 5.

1. a) 100 c) -98 e) 1

6. a) 12

7. a) 29

8. $\dfrac{131}{6}$

Ejercicios 6.

2. $\det A = a_{11}\, a_{22}\, a_{33}\, a_{44}$

4. No. Tome, por ejemplo, $A = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$; aquí, $\det A = \det B$ pero $A \neq B$.

5. Si $A^2 = A$ entonces $\det A^2 = \det A$ luego $\det A(\det A - 1) = 0$, y así $\det A = 0$ ó $\det A = 1$.

6. No. Tome por ejemplo $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$; así, aunque $A^T = -A$ se tiene que $\det A \neq 0$.

11. *a*) $X = t(\frac{5}{4}, 1) + (\frac{1}{2}, 0)$. *b*) $X = t(\frac{1}{4}, -\frac{13}{8}, 1) + (\frac{3}{2}, \frac{19}{4}, 0)$.

Ejercicios complementarios

1. $AB = \begin{bmatrix} 30 & 2 & 7 \\ 9 & -2 & -3 \\ 2 & 2 & 3 \end{bmatrix}$

2. $AB = 19, \quad BA = \begin{bmatrix} 21 & -7 \\ 6 & -2 \end{bmatrix}$

4. a) $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

   b) $B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

6. a) Debemos probar que $B^T = B$, y, en efecto:
   $B^T = (A + A^T)^T = A^T + (A^T)^T = A^T + A = A + A^T = B$

   b) $C^T = (A - A^T)^T = A^T - (A^T)^T = A^T - A = -(A - A^T) = -C$

   c) Tome $B = \frac{1}{2}(A + A^T)$ y $C = \frac{1}{2}(A - A^T)$.

14. 7

15. 3

26. a) 13

Lección 3. Matriz inversa

Ejercicios 1.

2. $(A^T)^{-1} = \begin{bmatrix} \frac{3}{7} & -\frac{1}{7} \\ \frac{1}{7} & \frac{2}{7} \end{bmatrix}$

5. La matriz

$$A = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$$

satisface $A^2 = 0$, y no puede ser inversible, porque si $A^2 = 0$ y A es inversible entonces, tomando inversa a ambos lados de la igualdad, llegaríamos a que $A = 0$. Es decir, no existe matriz inversible que satisfaga la condición $A^2 = 0$.

6. Tome cualquier dos matrices que no conmuten, es decir, que $AB \neq BA$. Para que la igualdad se cumpla es suficiente que las matrices conmuten.

8. b) $A^n = \begin{bmatrix} \cos(n\theta) & \operatorname{sen}(n\theta) \\ -\operatorname{sen}(n\theta) & \cos(n\theta) \end{bmatrix}$

8. c) Es inversible para cualquier valor de θ. La inversa de A es

$$A^{-1} = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

9. Debe darse que $(A+B)(A^{-1}+B^{-1}) = I_n$, es decir, que $AB^{-1}+BA^{-1} = -I_n$.

10. $(A^n)^{-1} = (A^{-1})^n$

Ejercicios 2.

1. a) $\begin{bmatrix} 2 & -3 & -2 \\ -1 & 3 & 2 \\ 0 & 1 & 1 \end{bmatrix}$

1. b) $\begin{bmatrix} -\frac{1}{19} & \frac{7}{19} & -1 \\ \frac{7}{19} & -\frac{11}{19} & 2 \\ -\frac{3}{19} & \frac{2}{19} & 0 \end{bmatrix}$

1. c) $\begin{bmatrix} \frac{7}{6} & -\frac{5}{3} & \frac{1}{6} \\ -\frac{1}{6} & \frac{2}{3} & -\frac{1}{6} \\ -1 & 2 & 0 \end{bmatrix}$

1. d) No existe.

2. a) Si $p = 0$ y $q = -1$, el sistema tiene infinitas soluciones.

2. b) Si $p = 0$ y $q \neq -1$, el sistema no tiene solución.

2. c) Si $p \neq 0$, el sistema tiene solución única:

$$x = \frac{2p + pq - q - 1}{3p}, \quad y = \frac{2q + pq - p + 2}{3p}, \quad z = -\frac{1+q}{p}$$

Ejercicios 3.

1 .a) $A^{-1} = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}$, $X = [-37, 22]$

1. d) $A^{-1} = \begin{bmatrix} \frac{1}{147} & \frac{3}{49} \\ \frac{5}{49} & -\frac{4}{49} \end{bmatrix}$, $X = [-\frac{2}{147}, \frac{39}{49}]$

2. a) $\begin{bmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{bmatrix}$

2. d) $\begin{bmatrix} \frac{1}{10} & 0 & 0 \\ 0 & -\frac{9}{23} & \frac{8}{23} \\ 0 & \frac{4}{23} & -\frac{1}{23} \end{bmatrix}$

3. $x_1 = \dfrac{21}{29}, \quad x_2 = \dfrac{171}{29}, \quad x_3 = -\dfrac{284}{29}, \quad x_4 = -\dfrac{182}{29}$

Ejercicios complementarios

1. $\det A = \det(P^{-1}BP) = \det P^{-1} \det B \det P = \dfrac{1}{\det P} \det B \det P = \det B$.

2. $A^{-1}(A+B)B^{-1} = (I + A^{-1}B)B^{-1} = B^{-1} + A^{-1}$. Así, obtenemos una fórmula para la inversa de una suma de dos matrices, en términos de las inversas de éstas:

$$(A+B)^{-1} = B^{-1}(A^{-1} + B^{-1})^{-1}A^{-1}$$

6. Si $\lambda_1, \lambda_2, ..., \lambda_n$ son las entradas de la diagonal principal de la matriz, entonces la matriz que tiene su diagonal conformada (en orden) por los recíprocos de estas entradas, es decir, por $\dfrac{1}{\lambda_1}, \dfrac{1}{\lambda_2}, ..., \dfrac{1}{\lambda_n}$, es la matriz inversa.

9. Aquellos a y b para los cuales $\det A = 0$, es decir, para los cuales $2ab - 6a - 3b + 24 = 0$.

12. Porque es triangular superior y tiene un 0 en la diagonal principal.

13. $(A^{-1})^T = (A^T)^{-1} = A^{-1}$

17. 40 cm^2

18. a) $p_1 = \frac{35}{136}p_3; \quad p_2 = \frac{55}{136}p_3$

18. b) $[\frac{11}{20}, \frac{103}{160}, \frac{11}{16}]$

Lección 4. Vectores

Ejercicios 1.

1. a) $(0,5)$ 1. d) $(-6,13)$ 1. e) $(0,-5)$

2. c) $(-17,-9,18)$ 2. d) $(25,-1,2)$

5. Los vectores en a), b), y f) son paralelos.

Ejercicios 2.

1. a) 2 1. d) $2\sqrt{5}$ 1. e) $3\sqrt{6}$

2. a) $\sqrt{13}$ 2. d) $\sqrt{34}$

3. La distancia entre $(1,1)$ y $(5,4)$ es 5, lo mismo que la distancia entre $(1,1)$ y $(-2,5)$.

5. Basta comprobarlo utilizando el teorema de Pitágoras: la distancia al cuadrado entre $(2,-1)$ y $(-1,-3)$ es 13; la distancia entre $(-2,5)$ y $(2,-1)$ es 52; la distancia al cuadrado entre $(-1,-3)$ y $(-2,5)$ es 65. Notamos que $13+52=65$.

Ejercicios 3.

1. a) 1 1. c) -5

2. En general, es falso. Tome, por ejemplo, $x=(1,0)$, $y=(1,2)$, $z=(1,3)$.

4. $\alpha = \dfrac{10}{3}$

7. a) Es el mismo $y=(1,0)$.

7. d) $\dfrac{15}{26}(3,1,4)$.

8. $P=A-B$ donde $A=(-\cos 30°, -\operatorname{sen} 30°)$ y $B=(-\cos 60°, \operatorname{sen} 60°)$. Este es un problema típico de equilibrio estático de fuerzas físicas.

9. a) $7i+14j-14k$

Ejercicios 4.

1. a) $X(x,y) = P+tN = (1,2)+t(3,4)$, es decir, $x = 1+3t$ y $y = 2+4t$ (ecuaciones paramétricas). La ecuación cartesiana se encuentra "despejando" t allí: $\frac{x-1}{3} = \frac{y-2}{4}$, ó, $4x - 3y + 2 = 0$.

2. a) Tomemos $P = (2,1,1)$ y $N = (3,2,6) - (2,1,1) = (1,1,5)$; entonces $X = P + tN = (2,1,1) + t(1,1,5)$, y así, con $X = (x,y,z)$, podremos escribir las ecuaciones paramétricas como $x = 2+t$, $y = 1+t$, $z = 1+5t$.

2. c) Tomando $P = (0,0,0)$ y $N = (6,-2,1)$ (este último es el vector normal al plano) tendremos $X = P + tN = t(6,-2,1)$, y así, con $X = (x,y,z)$, podremos escribir las ecuaciones paramétricas como $x = 6t$, $y = -2t$, $z = t$.

3. b) En la ecuación $(X - P)N = 0$, coloque $X = (x,y,z)$, $P = (1,2,2)$ y $N = (-1,2,-8)$.

3. c) En la ecuación $(X - P)N = 0$, coloque $X = (x,y,z)$, $P = (1,2,1)$ y $N = (1,1,1)$ (este último es ortogonal a $(3,2,-5)$ que es un vector ortogonal al plano $3x + 2y - 5z = 0$).

Ejercicios complementarios

1. a) $xy = 15$, $yz = -14$.

1. c) $(10,-1,6)$

3. $F_1 = \|F_1\|(-\cos 30°, \operatorname{sen} 30°)$, $F_2 = \|F_2\|(\cos 45°, \operatorname{sen} 45°)$ y $N = (0,50)$. De la igualdad de equilibrio estático de fuerzas $F_1 + F_2 = N$ se obtiene que $\|F_1\| \cos 30° = \|F_2\| \cos 45°$ y $\|F_1\| \operatorname{sen} 30° + \|F_2\| \cos 45° = 50$.

4. 0.5232 radianes.

7. En particular, si $x = y$ entonces tendríamos $\|x\|^2 = 0$, y así $x = 0$. Se ha probado que el único vector que es ortogonal a *todos* los vectores, es el vector nulo.

13. $c = -5$

15. La ecuación de la recta que pasa por $(1,2,3)y(9,2,0)$ es $X = (1,2,3) + t(8,0,-3)$. El punto $(2,8,1)$ no satisface esta ecuación para ningún t.

17. $x - y$ es ortogonal a $x + y$ si, y sólo si, $(x - y)(x + y) = \|x\|^2 - \|y\|^2 = 0$, lo que es equivalente a que $\|x\| = \|y\|$. La interpretación geométrica es que $x - y$ y $x + y$ son ortogonales si, y sólo si, x y y están en un mismo círculo.

19. La ecuación de la recta es $X = \frac{1}{2}(9, -1, 0) + t(-13, -1, 4)$

22. Como (l, m, n) es un vector paralelo a la recta (vector que dirige la recta) y (A, B, C) es un vector normal al plano, entonces éstos deben ser ortogonales, es decir, $Al + Bm + Cn = 0$.

24. a) $(-2, 1, 5)$

35. $d = \dfrac{|D|}{\sqrt{A^2 + B^2 + C^2}}$

Lección 5. Bases y dimensión

Ejercicios 1.

1. a) No es subespacio de $\mathbb{R}^n$ pues si a es un vector en el conjunto, $(-1)a = -a$ no lo es.

1. b) Es subespacio de $\mathbb{R}^n$.

1. c) Este conjunto no es subespacio de $\mathbb{R}^n$ pues 0 no es un vector en él.

1. d) Es subespacio de $\mathbb{R}^n$.

2. a) Es subespacio de $\mathcal{F}_{\mathbb{R}}$.

2. b) También es subespacio de $\mathcal{F}_{\mathbb{R}}$.

2. c) Es subespacio de $\mathcal{F}_{\mathbb{R}}$.

2. d) También es subespacio de $\mathcal{F}_{\mathbb{R}}$.

3. a) Es subespacio de $\mathfrak{M}_{n\times n}$.

3. b) También es subespacio de $\mathfrak{M}_{n\times n}$.

4. a) $\{t(1, 2, 3)/\ t \in \mathbb{R}\}$.

4. b) Todos los polinomios $p(x)$ de la forma $p(x) = (3a + b) + b\,x + a\,x^2 + a\,x^3$ para $a, b \in \mathbb{R}$ cualquiera.

4. d) $\{(0, 0)\}$

4. f) $\mathbb{R}^2$ 4. g) $\mathbb{R}^3$ 4. h) $\mathbb{R}^3$

5. a) Porque la combinación lineal de dos vectores, uno en un espacio vectorial, y el otro vector en el otro espacio vectorial, no necesariamente está en la unión de los dos espacios vectoriales. El problema está en que la operación (lineal) de tomar combinaciones es de naturaleza distinta a la de tomar la operación (conjuntista) de la unión. Por ejemplo, en una gráfica, tome la unión de las rectas $y = 3x$ y $y = 2x$, y convénzase de que la unión de ellas dos no contiene sus posibles combinaciones lineales.

5. b) El producto cartesiano de dos espacios vectoriales sí es, efectivamente, un espacio vectorial.

7. $(1, 2, 4) = 2(2, 0, 1) + (-3, 2, 2)$

8. Escriba primero $x^3 + 2x + 1 = \lambda_1 x + \lambda_2(3x + x^2) + \lambda_3(x^3 + x + 2)$ para ciertos números reales $\lambda_1, \lambda_2, \lambda_3$, y note, comparando el polinomio de la izquierda de la igualdad con el polinomio de la derecha, que $\lambda_1 = 1$, $\lambda_2 = 0$, $\lambda_3 = 1$.

Ejercicios 2.

2. $\alpha = \dfrac{24}{5}$

5. Ningún grupo de a tres vectores conforma una base para $\mathbb{R}^3$.

6. Basta tomar dos vectores linealmente independientes que estén en el plano $3x + y - z = 0$. Por ejemplo, $(1, 1, 4)$ y $(0, 1, 1)$.

Ejercicios 3.

3. Una base ortonormal es $\{\, (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}),\ (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})\, \}$.

5. Una base ortonormal es $\{\, (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0),\ (\frac{4}{\sqrt{57}}, \frac{4}{\sqrt{57}}, -\frac{5}{\sqrt{57}}),$
$(\frac{5}{\sqrt{114}}, \frac{5}{\sqrt{114}}, \frac{8}{\sqrt{114}})\, \}$.

6. a) $\{\, (0, \frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}}),\ (\frac{5}{\sqrt{70}}, -\frac{6}{\sqrt{70}}, -\frac{3}{\sqrt{70}})\, \}$

6. b) $\{\, (\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}),\ (0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})\, \}$

Ejercicios 4.

1. a) Dimensión 0. 1. d) Dimensión 0. 1. e) Dimensión 1.

Ejercicios complementarios

1. a) Es espacio vectorial.

1. b) También es espacio vectorial.

1. e) No es espacio vectorial pues 0 no está en el conjunto.

2. No es subespacio de $\mathbb{R}^n$ pues 0 no está en el conjunto.

3. a) No es subespacio de $\mathcal{F}_{\mathbb{R}}$ pues la multiplicación por escalar de una de estas funciones nos puede llevar a una función lineal que no tiene coeficientes enteros.

3. b) Es un subespacio de $\mathcal{F}_{\mathbb{R}}$.

4. a) No es subespacio de $\mathfrak{M}_{n\times n}$.

4. b) No es subespacio de $\mathfrak{M}_{n\times n}$.

4. c) Es subespacio de $\mathfrak{M}_{n\times n}$.

5. b) Es la recta $y = x$ en el plano $\mathbb{R}^2$.

5. c) Es todo el plano $\mathbb{R}^2$.

5. d) Es el plano $2x - y - z = 0$ en el espacio $\mathbb{R}^3$.

5. e) Son todas las funciones en $\mathcal{F}_{\mathbb{R}}$ que tienen la forma $f(x) = a + ax + be^x$ donde $a, b \in \mathbb{R}$.

6. $e^x + x + 1 = \dfrac{1}{2}f_1(x) - \dfrac{1}{2}f_2(x) + \dfrac{1}{2}f_3(x)$

7. a) $A = 2A_1 - A_2 - 2A_3$

7. b) No existen escalares c_1, c_2, c_3 tales que $A = c_1A_1 + c_2A_2 + c_3A_3$.

8. a) Una base puede ser $\{(1, -1, 0, 0), (0, 1, -1, 0), (0, 0, 1, -1)\}$. La dimensión es 3.

9. La dimensión es 4.

10. Una base puede ser $\{(1, 7, -5)\}$; la dimensión es 1.

12. a) El eje y.

12. c) El mismo plano que en 12.b).

15. Una base puede ser $\{(1,0,0),(0,1,1)\}$, y otra puede ser $\{(2,0,0),$ $(0,\frac{1}{2},\frac{1}{2})\}$.

16. No es cierto: en $\mathbb{R}^2$ se tiene como base a $\{(1,0),(0,1)\}$, pero el subespacio generado por el vector $(1,1)$ (que es la recta $y=x$ en el plano) no tiene a ninguno de los primeros vectores como base.

18. Dado el resultado $dim(W_1+W_2)=dim(W_1)+dim(W_2)-dim(W_1\cap W_2)$, y sabiendo que $dim(W_1)=3$, $dim(W_2)=3$, y $dim(W_1+W_2)\leq 5$, se tendrá que $dim(W_1\cap W_2)>0$.

20. Supongamos que $\lambda_1 x+\lambda_2(x+y)+\lambda_3(x+y+z)=0$; entonces $(\lambda_1+\lambda_2+\lambda_3)x+(\lambda_2+\lambda_3)y+\lambda_3 z=0$ y, por la independencia lineal de x,y,z, se tendrá que $\lambda_1+\lambda_2+\lambda_3=0$, $\lambda_2+\lambda_3=0$ y $\lambda_3=0$. De estas tres ecuaciones resulta que $\lambda_1=\lambda_2=\lambda_3=0$, lo que muestra que x, $x+y$ y $x+y+z$ son linealmente independientes.

21. Puesto que la fórmula general afirma que $x\cdot y=\|x\|\|y\|\cos\angle(x,y)$, entonces $x\cdot y=\|x\|\|y\|$ si, y sólo si, $\cos\angle(x,y)=1$; y esto es equivalente a que x y y son colineales, es decir, a que son linealmente dependientes.

26. El vector $(1,-1,1)$ conforma una base.

27. $\overline{X}=(\frac{215}{3},\frac{205}{3})$

28. a) $\overline{X}=(I_3-A)^{-1}C=(57.12,24.34,18.15)$

28. b) $p=w(I_3-A^T)^{-1}A_0=w(1.36,4.99,3.91)$

28. c) $x_{01}=a_{01}\overline{x_1}=57.12$; $\quad x_{02}=a_{02}\overline{x_2}=48.68$; $\quad x_{03}=a_{03}\overline{x_3}=54.45$

Lección 6. Transformaciones lineales

Ejercicios 1.

1. a) No es una transformación lineal pues $T(0,0)=(3,0)\neq(0,0)$.

1. d) No es una transformación lineal pues $T(1,1)=(1,1)$ pero $(4,2)=T(2,2)\neq 2T(1,1)=(2,2)$.

1. e) No es una transformación lineal pues $T(0,0)=(0,1)\neq(0,0)$.

2. $T(x, y, z) = (-y + z, \frac{1}{2}x + \frac{3}{2}y + z)$

3. Tome $A = \begin{bmatrix} 1 & -1 & 1 \\ -3 & 1 & 0 \\ -1 & -1 & -5 \end{bmatrix}$

5. Una posible interpretación geométrica en el caso bidimensional es que los nuevos ejes coordenados son simétricos con respecto a la recta $x = y$. En el caso tridimensional la simetría es con respecto a la recta $x = y = z$.

6. $\dfrac{x^2}{81} + \dfrac{y^2}{144} = 1$

7. a) Se convierte en la recta $BP + tBA$ donde $B = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$.

8. a) Se convierte en $A(X - P)(AN) = 0$ donde $A = \begin{bmatrix} 3 & -2 & 1 \\ 2 & -1 & 2 \\ 1 & 0 & -3 \end{bmatrix}$.

9. a) Basta notar que $AA^T \neq I_2$, para asegurar que $T(x) = Ax$ no es una transformación simétrica.

9. b) Basta notar que $AA^T = I_2$, para asegurar que $T(x) = Ax$ sí es una transformación simétrica.

10. $a = \dfrac{1}{\sqrt{6}}, \quad b = -\sqrt{\dfrac{2}{3}}, \quad c = \dfrac{1}{\sqrt{6}}$.

Ejercicios 2.

1. b) $Nu(T) = \{(0, 0, 0)\}; \quad Im(T) = \langle (1, 2, 1), (3, -1, 4), (-1, -1, 1) \rangle = \mathbb{R}^3$

1. c) El rango es 3, y la nulidad es 0.

2. a) i) El rango es 1.

2. a) ii) El rango es 2.

2. c) iii) El rango es 3.

Ejercicios 3.

1. a) $A = \begin{bmatrix} 1 & 1 \\ 2 & -3 \end{bmatrix}$

1. c) $A = \begin{bmatrix} 2 & 4 & 1 \\ 1 & 1 & -1 \\ 4 & -4 & 3 \end{bmatrix}$

2. a) $A = \begin{bmatrix} 1 & 1 \\ 1 & -2 \\ 4 & 3 \end{bmatrix}$

2. b) $A = \begin{bmatrix} 5 & 4 \\ -1 & -5 \\ 18 & 13 \end{bmatrix}$

4.a) $B = P^{-1}AP$ donde $P = \begin{bmatrix} -2 & 1 \\ 3 & 1 \end{bmatrix}$

4. b) $B = P^{-1}AP$ donde $P = \begin{bmatrix} 2 & -1 & -1 \\ 1 & 0 & 2 \\ 2 & 1 & 0 \end{bmatrix}$

5. Puesto que A tiene rango completo si, y sólo si, $\rho(A) = \text{mín}\, m, n = n$ entonces el resultado es inmediato ya que $\rho(A^T A) = \rho(A)$.

6. a) El rango fila y el rango columna es 3. Tiene rango completo.

6. b) El rango fila y el rango columna es 2. Tiene rango completo.

8. Note que la tercera fila de la matriz es combinación lineal de las dos primeras: $6(3,0,2,2) - \dfrac{1}{2}(-6,42,24,54) = (21,-21,0,-15)$.

9. a) El núcleo es $\langle (2,3,5) \rangle$ y tiene dimensión 1.

9. b) La dimensión de la imagen de T es $dim(\mathbb{R}^3) - dim(Nu(T)) = 3 - 1 = 2$.

9. c) El rango de A es 2.

9. d) El rango de A coincide con la dimensión de la transformación T.

Ejercicios 4.

2. Es la rotación del sistema coordenado estándar por un ángulo $n\theta$.

Ejercicios 5.

2. $T(\lambda A + \beta B) = P^{-1}(\lambda A + \beta B)P = \lambda P^{-1}AP + \beta P^{-1}BP = \lambda T(A) + \beta T(B)$ para escalares cualquiera λ, β. Esta transformación lineal es la representación algebraica del problema geométrico de primero encontrar las coordenadas de un vector bajo un diferente sistema coordenado, y luego retornar el vector a su representación en el sistema coordenado original.

Ejercicios complementarios

1. a) Es transformación lineal.

1. b) No es transformación lineal pues $T(0,0) \neq (0,0)$.

2. No es transformación lineal porque $\det(A+B) \neq \det A + \det B$.

6. a) $\rho(T) = 3$ 6. b) $\rho(T) = 3$

7. a) $\begin{bmatrix} 1 & -1 & 2 & -1 \\ -1 & 0 & 2 & 1 \\ 2 & 2 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$

7. b) $\begin{bmatrix} -34 & -\frac{11}{2} & \frac{175}{2} & 5 \\ -\frac{83}{2} & -11 & \frac{217}{2} & \frac{3}{2} \\ -\frac{33}{2} & -3 & \frac{87}{2} & \frac{5}{2} \\ \frac{65}{2} & \frac{19}{2} & -84 & -\frac{1}{2} \end{bmatrix}$

7. c) $P = \begin{bmatrix} 1 & 3 & 0 & 5 \\ -2 & 0 & 4 & 0 \\ 1 & 2 & -1 & 3 \\ 1 & -2 & 1 & -1 \end{bmatrix}$

8. a) $\rho(T) = 1,\ nul(T) = 1;\quad \dim \mathbb{R}^2 = \rho(T) + nul(T)$

8. c) $\rho(T) = 2,\ nul(T) = 1;\quad \dim \mathbb{R}^3 = \rho(T) + nul(T)$

Lección 7. Diagonalización en $\mathbb{R}^n$

Ejercicios 1.

1. a) Los valores propios son $\lambda = 1 + 2\sqrt{2}$ con vector propio $(2(\sqrt{2}+1), 1)$, y $\lambda = 1 - 2\sqrt{2}$ con vector propio $(2(1-\sqrt{2}), 1)$,

1. c) Los valores propios son $\lambda = 1 + 2\sqrt{5}$ con vector propio $(1, 3+\sqrt{5})$, y $\lambda = 1 - 2\sqrt{2}$ con vector propio $(2(1-\sqrt{2}), 1)$.

1. e) Los valores propios son $\lambda = 1$ con vector propio $(0,1,0)$, y $\lambda = -3$ con vector propio $(-4,1,2)$.

2. a) Los valores propios de una matriz diagonal son, precisamente, las entradas de la diagonal principal. Así, el único valor propio de la matriz I_n es 1 con multiplicidad n.

2. b) Puesto que $\det(kA - k\lambda I_n) = k^n \det(A - \lambda I_n) = 0$, los valores propios de kA son de la forma $k\lambda$ para λ un valor propio de A.

Ejercicios 2.

1. $P^{-1}AP = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

3. Sí es diagonalizable, y los valores propios son $\lambda = 2$ con vector propio $[1, 8, -10]$; $\lambda = \sqrt{39}$ con vector propio $[1, 1, \frac{3-\sqrt{39}}{2}]$; y $\lambda = -\sqrt{39}$ con vector propio $[1, 1, \frac{3+\sqrt{39}}{2}]$.

4. b) Esta matriz tiene un único valor propio $\lambda = 2$ con multiplicidad 3, y con espacio propio asociado $\langle(1, 0, 0)\rangle$.

5.c) Los valores propios son $\lambda_1 = 1$ con espacio propio asociado $\langle(1, 0, 0, 0)\rangle$; $\lambda_2 = -1$, con espacio propio asociado $\langle(0, 1, 0, 0)\rangle$; y $\lambda_3 = \lambda_4 = 2$ con espacio propio asociado $\langle\{(0, 0, 1, 0), (1, -\frac{2}{5}, 0, -\frac{1}{5})\}\rangle$.

6. a) Puesto que $\det(A - \lambda I_n) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$ entonces, haciendo aquí $\lambda = 0$, se tiene

$$\det A = (-1)^n(-\lambda_1)(-\lambda_2) \cdots (-\lambda_n) = \lambda_1\lambda_2 \cdots \lambda_n$$

8. Supongamos que A es inversible. Entonces tendremos que $\det(BA - \lambda I) = \dfrac{1}{\det A}\det(A(BA) - \lambda A) = \dfrac{1}{\det A}\det((AB)A - \lambda A) = \det(AB - \lambda I)$.

Ejercicios 3.

1. a) i) Los valores propios son $\lambda_1 = \frac{3+\sqrt{37}}{2}$ con espacio propio $\langle(6, 1 + \sqrt{37})\rangle$; y $\lambda_2 = \frac{3-\sqrt{37}}{2}$ con espacio propio $\langle(6, 1 - \sqrt{37})\rangle$.

1. a) ii) $P = \begin{bmatrix} \frac{6}{\sqrt{74+2\sqrt{37}}} & \frac{6}{\sqrt{74-2\sqrt{37}}} \\ \frac{1+\sqrt{37}}{\sqrt{74+2\sqrt{37}}} & \frac{1-\sqrt{37}}{\sqrt{74-2\sqrt{37}}} \end{bmatrix}$

1. c) i) Los valores propios son $\lambda_1 = 6$ con espacio propio $\langle(-1, 5, 3)\rangle$; y $\lambda_2 = -1$ con espacio propio $\langle(-1, -2, 3)\rangle$; $\lambda_3 = 1$ con espacio propio $\langle(3, 0, 1)\rangle$.

1. c) ii) $P = \begin{bmatrix} -\frac{1}{\sqrt{35}} & -\frac{1}{\sqrt{14}} & \frac{3}{\sqrt{10}} \\ \frac{5}{\sqrt{35}} & -\frac{2}{\sqrt{14}} & 0 \\ \frac{3}{\sqrt{35}} & \frac{3}{\sqrt{14}} & \frac{1}{\sqrt{10}} \end{bmatrix}$

Ejercicios 4.

1. a) $Q = XAX^T$ donde $A = \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}$. Esta forma cuadrática no es de ninguna de las clases indicadas pues los valores propios son 4 y -1.

1. c) $Q = XAX^T$ donde $A = \begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix}$. Esta forma cuadrática es semidefinida negativa pues los valores propios son 0 y -10.

2. a) $Q = XAX^T$ donde $A = \begin{bmatrix} 3 & 7 \\ 7 & -6 \end{bmatrix}$. Los valores propios son $\frac{\sqrt{277}-3}{2}$ y $\frac{-\sqrt{277}-3}{2}$. La cuadrática se puede entonces transformar linealmente en la hipérbola

$$(\frac{-\sqrt{277}-3}{2})(y_1)^2 + (\frac{\sqrt{277}-3}{2})(y_2)^2 = 5$$

2. c) $Q = XAX^T$ donde $A = \begin{bmatrix} 3 & 2\sqrt{3} \\ 2\sqrt{3} & 7 \end{bmatrix}$. Los valores propios son 9 y 1. La cuadrática se puede entonces transformar linealmente en la elipse

$$9(y_1)^2 + (y_2)^2 = 9$$

Ejercicios 5.

1. b) La matriz Jordan es $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

1. c) La matriz Jordan es $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 4+\sqrt{11} & 0 \\ 0 & 0 & 4-\sqrt{11} \end{bmatrix}$

Ejercicios complementarios

1. Los valores propios son $\lambda = \pm i$. Representa una rotación de 90°. (Para comprender el significado de los valores propios complejos, ver el ejercicio 17 de estos ejercicios complementarios.)

3. $P^{-1}AP = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} d & c \\ b & a \end{bmatrix}$

5. Los valores propios de esta matriz son $\lambda_1 = \lambda_2 = \lambda_3 = \lambda = 0$ para todo $a, b, c \in \mathbb{R}$. Y aunque $\lambda = 0$ tiene multiplicidad 4, su espacio propio sólo tiene dimensión 1.

7.) $P^{-1}AP = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

9. a) La ecuación característica es $(2-\lambda)^3 = 0$, y el único vector propio asociado a $\lambda = 2$ es $(0, 0, 0)$.

14. a) $P^T = (A(A^TA)^{-1}A^T)^T = A((A^TA)^{-1})^TA^T = A((A^TA)^T)^{-1}A^T = A(A^TA)^{-1}A^T = P$

16. La matriz Jordan es $\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

Lección 8. Conjuntos convexos

Ejercicios 1.

3. Consideremos el hiperplano H cuya ecuación cartesiana es $px = b$. Sean, además, x_1 y x_2 dos puntos arbitrarios de H. Entonces, inmediatamente notamos que, para $\lambda \in [0, 1]$, también el vector $\lambda x_1 + (1-\lambda)x_2$ está en H, pues

$$p(\lambda x_1 + (1-\lambda)x_2) = \lambda p x_1 + (1-\lambda)p x_2 = \lambda b + (1-\lambda)b = b$$

4. Puesto que si $\|x\| = 1$ entonces $\|-x\| = 1$, tendremos que, para $\lambda \in [0, 1]$,

$$\|\lambda x + (1-\lambda)(-x)\| = \|(2\lambda - 1)x\| = |2\lambda - 1|$$

y $|2\lambda - 1| \neq 1$ si $\lambda \neq 1,\ 0$. La envolvente convexa es el conjunto $\{x \in \mathbb{R}^n / \ \|x\| \leq 1\}$.

6. a) $C_1 + C_2 = \{(x, y) \in \mathbb{R}^2 / (x-3)^2 + (y-4)^2 \leq 1\}$. En efecto: cualquier $a = b + (3, 4)$ con $b \in C_1$, satisface $\|a - (3, 4)\| = \|(b + (3, 4)) - (3, 4)\| \leq 1$; es decir, a pertenece al círculo de radio 1 con centro en $(3, 4)$.

6. c) $C_1 + C_2 = [2, 4] \times [2, 4]$

7. Si W es un espacio vectorial en $\mathbb{R}^n$, y $x,\ y \in W$ entonces, por definición de W, también $\lambda\, x + (1-\lambda)y \in W$ para $\lambda \in [0, 1]$.

8. Sólo probaremos la parte "solo si" del teorema, es decir, que si C es un cono convexo (que, sin pérdida de generalidad, asumiremos con vértice en el origen) entonces cualquier combinación lineal no-negativa de puntos de C es, nuevamente, un punto de C. En efecto: si tomamos dos puntos x, y en el cono C, y dos números no-negativos a y b (que podemos suponer no simultáneamente nulos), entonces notando que

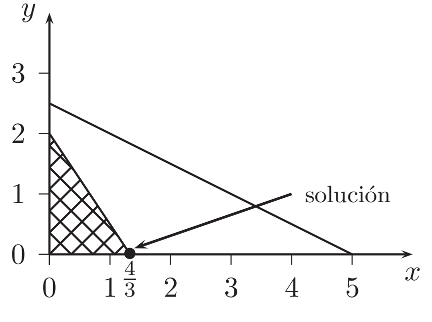
$$ax + by = (a+b)(\frac{a}{a+b}\,x + \frac{b}{a+b}\,y)$$

vemos que $ax+by$ es la multiplicación por el escalar $a+b$ de una combinación convexa de x y y, y esto demuestra la parte "solo si" del teorema.

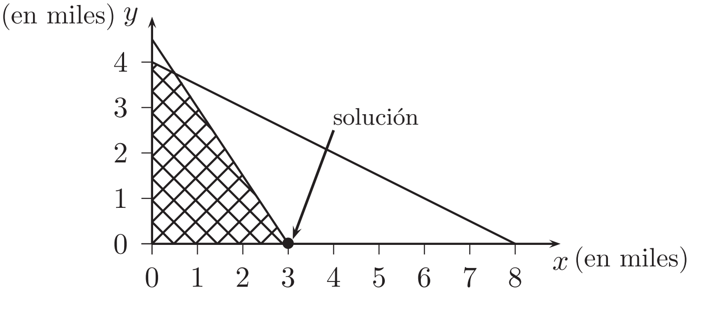
9. a) Este conjunto es un politopo que no es poliedro, pero que sí es un cono con vértice en $(0,0)$.

Ejercicios 2.

1. a)



b)



2. No tiene solución.

3. a) $s \geq 0$ y $r \geq 10$.
   b) $s \leq 0$.
   c) $r < -10$ y $10 < -rs$ con $s > 0$.

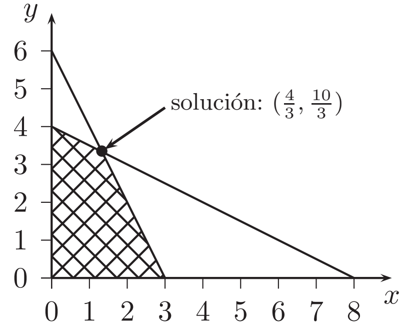
Ejercicios complementarios

1. a) Convexo

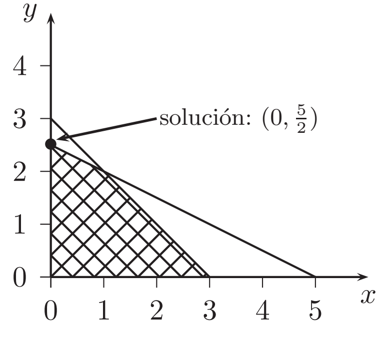
1. b) No convexo; el mínimo conjunto convexo que lo contiene es $\{(x, y) \in \mathbb{R}^2 / y \geq x^2\}$.

1. c) Convexo 1. f) Convexo 1. i) Convexo

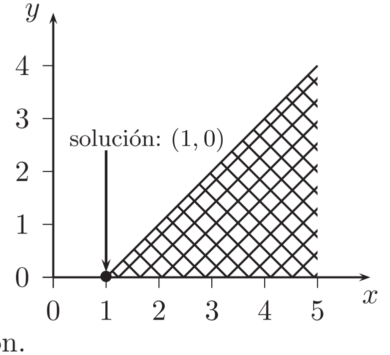
5. a)



b)



c)



d) No tiene solución.

7. Algunas soluciones enteras: probetas pequeñas: 8000; probetas grandes: 7000; probetas pequeñas: 8000, probetas grandes: 24000; probetas pequeñas: 10000, probetas grandes: 5000; probetas pequeñas: 9000, probetas grandes: 27000; probetas pequeñas: 26000, probetas grandes: 10000.

9. $x = 96, \;\; y = 0, \;\; z = 48.$

Índice alfabético

Álgebra lineal, 1
en la teoría económica, 18
Área
de un triángulo, 65
del paralelogramo, 63
Óptimo de Pareto, 381

Aditividad, 377
Al-Khwarizmi, 2
Algoritmo de eliminación gaussiano, 4
Análisis de actividades, 18, 21
Análisis insumo-producto, 18, 19, 219
Arrow, Kenneth, 19

Böhm-Bawerk, Eugen, 121
Babilonios, 2
Base de vectores propios, 300
Base(s), 196
canónica para $\mathbb{R}^n$, 201
ortogonal, 210
ortonormales, 209
para un espacio vectorial, 200
Biyectividad
y linealidad, 255
Bloque Jordan canónico, 320

Cassel, Gustave, 18–21, 166
Cauchy, Augustin L., 80, 157
Cayley, Arthur, 63, 157
Clifford, William K., 143
Combinación convexa, 348
Combinaciones lineales, 186
en $\mathbb{R}^4$, 188
en el espacio de funciones, 189
en el espacio de polinomios, 189
en $\mathbb{R}^2$, 186
en $\mathbb{R}^3$, 187
Conjuntos convexos, 341
propiedades básicas de, 343
Conos, 347
convexos, 347
Consumidores, 374
Convexidad en economía, 356
Costo de oportunidad, 121
Cournot, Augustin, 84
Cramer, Gabriel, 80, 111
Cuaterniones, 179
Curva IS lineal, 358
Curva LM lineal, 358

Dantzig, George B., 22, 353
Davidson, Paul, 361
Debreu, Gerard, 19
Dependencia lineal, 196
Desargues, Girard, 235
Descartes, René, 2, 131
Desigualdad Cauchy-Schwarz, 147, 148
Determinantes
2×2, 63
3×3, 64
$n \times n$, 68
del producto, 78
propiedad aditiva de, 75
propiedad de intercambio de filas de, 76
propiedad escalar de, 74
propiedades de los, 73
Diagonalización, 299
condición general para, de una matriz, 304
condición suficiente para, de una matriz, 302
de matrices simétricas, 307
de una matriz, 301
en bloques de Jordan, 319
Dimensión, 196

Diofanto, 2
Distancia a rectas y planos, 176
Distancia en $\mathbb{R}^n$, 140

Ecuación matricial, 93
Edgeworth, Francis, 20
Eliminación gaussiana, 12, 14, 15
Envolvente convexa, 348
Equilibrio competitivo, 380
Equilibrio de mercado, 379
Equilibrio físico, 138
Equilibrio walrasiano, 380
Escalar, 132
Espacio tridimensional, 135
Espacio vectorial, 180
 dimensión de un, 203
 soluciones del sistema homogéneo como, 183
Espacios vectoriales isomorfos, 277
Estrategia maxmin, 364
Estrategia minmax, 364
Estrategia mixta, 366
Euler, Leonhard, 2

Fermat, Pierre de, 131
Formas cuadráticas, 310
 definidas negativas, 314–316
 definidas positivas, 314, 315
 problema de reducción, 311
 semidefinidas negativas, 314, 318
 semidefinidas positivas, 314, 318
Función de bienestar económico, 22
Función de bienestar social, 22

Galilei, Galileo, 131
Gauss, Karl F., 2, 15, 80
Geometría
 de invariantes, 240
Geometría analítica, 131, 132
Geometría proyectiva, 235
Gram-Schmidt, 210
Grassmann, Hermann, 157

Hamilton, William, 157, 179
Hansen, Alvin, 357
Harrod, Roy, 357
Herón de Alejandría, 138
Hesse, Ludwig, 111
Hicks, John, 357
Hipérbola, 236
Hiperplano, 345
 afín, 345
Hiperplanos, 183

Imagen, 251
Independencia lineal, 196
Isomorfismos, 277

Jevons, William, 20, 121
Juego de matriz, 369
Juego de piedra-papel-tijera, 365
Juego de tirar la moneda, 365
Juegos de dos jugadores y de suma cero, 362

Kantorovich, Leonid, 353
Karush, William, 353
Kepler, Johannes, 131
Keynes, John Maynard, 18, 20
Koopmans, Tjalling, 19, 21, 22, 353

Laderman, Jack, 354
Lagrange, Joseph L., 80, 132
Lange, Oskar, 357, 373
Laplace, Pierre S., 80
Leibniz, Gottfried W., 80
Leontief, Wassily, 19, 219
Lerner, Abba, 358, 361
Ley de los cosenos, 145
Ley del paralelogramo, 136, 138
Longitud de un segmento, 65

Método gaussiano
 para el cálculo de la inversa, 100
Método simplex, 353
Marshall, Alfred, 20, 84
Marx, Karl, 18, 166
Matrices similares, 263
Matriz, 4, 30
 álgebra de inversas, 95
 adjunta, 111
 antisimétrica, 82
 aumentada, 4
 determinante de una, 63
 determinante de una, particionada, 116

diagonal, 32
elemental, 47
idéntica, 32
idempotente, 82
igualdad de, 31
inversa, 94
inversa de una, particionada, 117
invertible, 94, 116
Jordan canónica, 320
multiplicación, 38
multiplicación por escalar, 36
no-singular, 94
nula, 32
ortogonal, 213
particionada, 58
propiedades de la multiplicación, 42
propiedades de la traspuesta de una, 53
propiedades de multiplicación por escalar, 37
propiedades de suma, 37
propiedades de, simétricas, 55
propiedades del rango de una, 265
rango completo, 267
rango de una, 264
rotación, 99
simétrica, 55, 308
singular, 94
suma de, 35
traspuesta de una, 52
traza de una, 56
triangular, 33
Meade, James, 357
Menger, Carl, 20, 121
Modelo de análisis de actividades, 373
posibilidad de producción, 379
disponibilidad de recursos, 379
imposibilidad de producción, 379
Modelo de equilibrio general Walras-Cassel, 166
Modelo keynesiano IS-LM lineal, 357
equilibrio, 359
Modelo teórico de Sraffa, 323
Modelo Walras-Cassel, 18
Modelos de programación lineal, 22
Momento de una fuerza, 155
Monge, Gaspard, 157
Morgenstern, Oskar, 22, 23

Núcleo, 251
Números complejos, 179, 185
Neisser, Hans, 21

Operación fila, 8, 47
Operaciones fila
y multiplicación de matrices, 48
Operador lineal, 241
Ortogonalización Gram-Schmidt, 210

Parábola, 236
Pareto, Vilfredo, 20, 121
Pigou, Arthur, 20
Plano bidimensional, 134
Poliedros, 346
Polinomio característico, 295
Politopos, 346
acotado, 346
Poncelet, Jean Victor, 235
Precios de mercado, 121
Primer modelo lineal formal en la teoría económica, 84
Principio débil de bienestar, 22
Proceso de Gram-Schmidt, 210
Producto interior, 143
Productores, 375
Programación lineal, 350
Proyección de x sobre y, 151

Quesnay, François, 18

Recta real, 134
Recta(s) y plano(s), 157
en $\mathbb{R}^n$, 157, 161
Reducción a ejes principales, 312
Región factible o alcanzable, 351
Regla de Cramer, 108, 111, 114
cálculo de la matriz inversa por la, 108
Rendimientos constantes a escala, 378
Ricardo, David, 18, 166
Robbins, Lionel, 23

Samuelson, Paul, 358
Saturación local, 376
Schlesinger, Karl, 21

Schwartz, Hermann, 148
Scitovsky, Tibor, 361
Semiespacio cerrado inferior, 346
Semiespacio cerrado superior, 346
Semiespacios cerrados, 351
Sistema de ecuaciones lineales, 1–3
    homogéneo, 2
    no-homogéneo, 2
Sraffa, Piero, 20, 323
Subespacio(s)
    de $\mathbb{R}^3$, 191
    generado por vectores, 192
    vectoriales, 189

Teoría de juegos, 18, 22, 361
Teoría de la imputación, 121
Teorema de Brouwer, 281
Teorema del minimax, 369
Teorema del punto fijo de Kakutani, 281
Teorema espectral, 307
    para matrices simétricas, 308
Teorema fundamental del álgebra lineal, 252
Transformación
    afín, 235
    invertible, 276
    propiedades de la, inversa, 276
Transformación(es) lineal(es), 240
    composición de, 275
    de Lorentz, 245
    estructura de los conjuntos de, 272
    ortogonales, 247
    propiedades del producto de, 275
    suma y producto por escalar de, 272
    y matrices, 258

Valor
    de intercambio, 121
    de uso, 121
    intrínseco, 121
    teoría del, 121
Valor del juego, 364
Valores propios, 293
    de una transformación lineal, 293
Vandermonde, Alexandre T., 80, 90
Variable algebraica, 132
Vector(es), 132–134
    álgebra básica del producto cruz, 155
    álgebra de, 136
    ángulo entre, 143, 146
    de norma 1, 141
    igualdad de, 133, 135
    norma de un, 133
    norma de un, en $\mathbb{R}^n$, 139
    norma de un, y distancia en $\mathbb{R}^2$, 139
    norma de un, y distancia en $\mathbb{R}^3$, 140
    norma de un, y distancia en $\mathbb{R}^n$, 140
    ortogonal, 148
    ortogonal en $\mathbb{R}^2$, 149
    ortogonal en $\mathbb{R}^3$, 149
    paralelos, 137
    producto cruz, 151, 156
    producto interior, 143
    producto por escalar, 136
    producto punto entre, 143
    propiedades de la norma de un, 149
    propiedades del producto cruz, 153
    propiedades del producto interior, 144
    proyección de un, sobre otro, 150
    resta de, 136
    suma de, 136
    unitario, 141
Vectores ortogonales, 209
Vectores propios, 293
    de una transformación lineal, 293
Viete, François, 2
Volumen
    de un tetraedro, 71
    del paralelepípedo como un determinante, 67
von Hayek, Friedrich, 373
von Neumann, John, 20, 22, 23, 281, 282, 284, 285, 287, 291, 353
    modelo de equilibrio general de, 281
von Stackelberg, Heinrich, 21
von Wieser, Friedrich, 121

Wald, Abraham, 19, 21
Walras, Léon, 19, 20, 121, 166
Weintraub, Sidney, 361

Matemáticas básicas para economistas
Sergio Monsalve (editor)

Volumen 0: Fundamentos

Lección 1. Sobre la geometría, la aritmética y la trigonometría griegas.
Lección 2. El álgebra de los siglos XVI y XVII.
Lección 3. La geometría analítica de Descartes y Fermat.
Lección 4. Sobre los fundamentos para las matemáticas contemporáneas.

Volumen 1: Álgebra Lineal

Lección 1. Sistemas de ecuaciones lineales: solución por eliminación gaussiana.
Lección 2. Matrices y determinantes.
Lección 3. Sistemas de ecuaciones lineales: solución por matriz inversa.
Lección 4. Vectores.
Lección 5. Bases y dimensión.
Lección 6. Transformaciones lineales.
Lección 7. Diagonalización en $\mathbb{R}^n$.
Lección 8. Conjuntos convexos.

Volumen 2: Cálculo

Lección 1. El método de límites.
Lección 2. La derivada.
Lección 3. Elementos básicos de la teoría de la optimización.
Lección 4. La integral.

Volumen 3: Optimización y dinámica

Lección 1. Funciones cóncavas, convexas, cuasicóncavas y cuasiconvexas.
Lección 2. Optimización estática.
Lección 3. Sistemas dinámicos.
Lección 4. Optimización dinámica.

Este libro es el resultado de varios años de trabajo de los autores como profesores de matemáticas o economía para las facultades de Ciencias y de Ciencias Económicas de la Universidad Nacional de Colombia (sedes Medellín y Bogotá), la Universidad Externado de Colombia y la Pontificia Universidad Javeriana, y su objetivo central es exponer algunos de los elementos fundamentales del lenguaje matemático que deberían ser comunes a todos los estudiantes de economía de nuestras épocas. Pensando en esto, optamos por escribir el texto en cuatro volúmenes: en el volumen 0 (Fundamentos) presentamos los requisitos matemáticos que el estudiante debe llenar para acceder más cómodamente al corpus total; el volumen I consiste en las nociones básicas del álgebra lineal; el volumen II, en las nociones básicas del cálculo diferencial e integral, y el volumen III, en las nociones básicas de la teoría de la optimización y de la dinámica.

En cada uno de los cuatro volúmenes dividimos los temas en *lecciones* con un *tratamiento matemático riguroso y sin referencia a aplicación económica alguna*. Todas estas lecciones presentan, además, notas históricas que esperamos ayuden a trazar el devenir de los conceptos matemáticos que se desarrollan al punto. Por tanto, quienes consideran que un curso de matemáticas básicas para economistas debería ser solo eso y no un curso con aplicaciones, estarán aquí servidos. Sin embargo, para los que difieren de esta postura metodológica y pedagógica, *también separamos la sección final de casi todas las lecciones para el "contexto económico"*, que no es una sección ordinaria de aplicaciones a la economía: es, por el contrario, una aproximación coherente a problemas centrales en la teoría económica, y una orientación para el estudiante atento y disciplinado.

Sergio Monsalve
Editor

Profesor asociado del Departamento de Matemáticas, Universidad Nacional de Colombia. Especializado en teoría de juegos clásica y evolutiva y en equilibrio general walrasiano, su labor académica y profesional se ha caracterizado por una marcada preocupación por la formación de economistas con rigor e intuición, atentos al desarrollo y a las fuentes históricas de las ideas.

Entre sus publicaciones se cuentan *Introducción a los conceptos de equilibrio en economía* (1999, 2002), *Un curso de teoría de juegos clásica* (2005) y diversos artículos en revistas nacionales e internacionales.

MATEMÁTICAS BÁSICAS
PARA ECONOMISTAS: 1

ÁLGEBRA LINEAL

Una vez presentado el volumen 0 (*Fundamentos*) de la colección *Matemáticas básicas para economistas*, creímos que el primer paso en la formación matemática de todo economista moderno era afrontar el estudio de las herramientas que permiten abordar "problemas lineales", es decir, de lo que hoy llamamos *álgebra lineal*. Al plantearlo así, decidimos tomar, como hilo articulador, *la solución de un sistema de ecuaciones lineales*, pues este problema, aparentemente simple, es el verdadero origen de una gran cantidad de conceptos e ideas del álgebra lineal: matriz, determinante, base, dimensión, etc.